

84254

From: Chan, Christina
Sent: Tuesday, January 14, 2003 3:04 PM
T : Yu, Misook; STIC-Biotech/ChemLib
Subject: RE: rush search for 09/648,310

Importance: High

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

CRF

-----Original Message-----

From: Yu, Misook
Sent: Tuesday, January 14, 2003 11:43 AM
To: Chan, Christina
Subject: rush search for 09/648,310

Please approve rush searh. The case is due this bi-week.

Please search SEQ ID NO:1 and 2.

~~Examiner Misook YU, Ph.D.~~
703-308-2454 (Phone)
Art Unit 1642
CM1-8E18 (Room)
CM1-8E12 (Mail Box)

Point of Contact:
Barb O'Bryen
Technical Information Specialist
STIC CM1 6A05 308-4291

Searcher: 1003
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 1-17-03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

BioTech-Chem Library

Search Results

Feedback Form (Optional)



Scientific & Technical Information Center

The search results generated for your recent request are attached. If you have any questions or comments (compliments or complaints) about the scope or the results of the search, please contact *the BioTech-Chem searcher* who conducted the search *or contact*:

Mary Hale, Supervisor, 308-4258
CM-1 Room 1E01

Voluntary Results Feedback Form

➤ *I am an examiner in Workgroup:* (Example: 1610)

➤ *Relevant prior art found, search results used as follows:*

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ *Relevant prior art not found:*

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Search results were not useful in determining patentability or understanding the invention.

Other Comments:

Drop off completed forms at the Circulation Desk CM-1, or send to Mary Hale, CM1-1E01 or mary.hale@uspto.gov

Yu, Misook

To: Chan, Christina
Subject: rush search for 09/648,310

Please approve rush search. The case is due this bi-week.

Please search SEQ ID NO:1 and 2.

Examiner Misook Yu, Ph.D.
703-308-2454 (Phone)
Art Unit 1642
CM1-8E18 (Room)
CM1-8E12 (Mail Box)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 21:46:28 ; Search time 102 Seconds
(without alignments)
3409.581 Million cell updates/sec

Title: US-09-648-310-1

Perfect score: 780

Sequence: 1 ggcacgagctctctctgccc.....gtaataaaaaaaatcatgt 780

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	288.2	36.9	778	10	US-09-925-300-545
2	206.6	26.5	552	10	US-09-920-300A-939
3	206.6	26.5	552	12	US-10-033-528-939
4	122.4	15.7	388	10	US-09-925-299-318
5	109.6	14.1	486	9	US-09-938-842A-355
6	92.8	11.9	365	9	US-10-046-935-2101
7	92.8	11.9	365	9	US-09-878-178-2101
8	92.2	11.8	406	10	US-09-880-107-687
9	87.4	11.2	268	10	US-09-294-0938B-2020
10	63.2	8.1	289	10	US-09-294-0938B-4150
11	38.6	4.9	6222	9	US-10-098-841-105
12	38.4	4.9	4259	8	US-08-927-939-30
13	38	4.9	2000	9	US-09-938-842A-3235
14	37.6	4.8	338	9	US-10-046-935-342
15	37.6	4.8	338	9	US-09-878-178-342
16	37.6	4.8	1223	9	US-09-938-842A-3570
17	37.6	4.8	1967	10	US-09-925-301-565
18	37.6	4.8	2431	10	US-09-822-830A-395
19	37	4.7	2000	9	US-09-938-842A-4612

C 20	37	4.7	12286	10	US-09-813-742-1	Sequence 1, Appli
C 21	37	4.7	640681	10	US-09-790-988-1	Sequence 1, Appli
C 22	36.6	4.7	2000	9	US-09-938-842A-4732	Sequence 4732, Ap
C 23	36.2	4.6	243	10	US-09-969-373-912	Sequence 912, App
C 24	36.2	4.6	243	10	US-09-969-373-912	Sequence 913, App
C 25	36.2	4.6	248	9	US-09-754-853A-157	Sequence 157, App
C 26	36.2	4.6	248	9	US-09-754-853A-158	Sequence 158, App
C 27	36.2	4.6	1446	9	US-09-938-842A-3392	Sequence 3392, Ap
C 28	36.2	4.6	335913	9	US-09-754-853A-2	Sequence 2, Appli
C 29	36.2	4.6	335913	9	US-09-754-853A-3	Sequence 3, Appli
C 30	36	4.6	49984	10	US-09-739-457-5	Sequence 5, Appli
C 31	35.8	4.6	213	10	US-09-969-373-1157	Sequence 1157, Ap
C 32	35.8	4.6	775	9	US-10-015-219-605	Sequence 605, App
C 33	35.8	4.6	775	10	US-09-777-564-605	Sequence 605, App
C 34	35.6	4.6	575	10	US-09-864-761-20733	Sequence 20733, A
C 35	35.6	4.6	575	10	US-09-864-761-3972	Sequence 3972, Ap
C 36	35.4	4.5	1969	9	US-10-114-893-77	Sequence 77, Appl
C 37	35.4	4.5	915	9	US-10-114-893-314	Sequence 314, App
C 38	35.4	4.5	2000	9	US-09-938-842A-4264	Sequence 4264, A
C 39	35.2	4.5	375	10	US-09-960-352-15014	Sequence 15014, A
C 40	35.2	4.5	51309	9	US-09-754-853A-4	Sequence 4, Appli
C 41	35	4.5	4237	10	US-09-962-832-117	Sequence 117, App
C 42	34.8	4.5	309	10	US-09-867-701-131	Sequence 131, App
C 43	34.8	4.5	640681	10	US-09-790-988-1	Sequence 1, Appli
C 44	34.6	4.4	618	10	US-09-878-574-4572	Sequence 4572, Ap
C 45	34.6	4.4	2000	9	US-09-938-842A-3707	Sequence 3707, Ap

ALIGNMENTS

RESULT 1

US-09-925-300-545
; Sequence 545, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 545
; LENGTH: 778
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (641)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (652)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-545

Query Match 36.98; Score 288.2; DB 10; Length 778;
Best Local Similarity 69.98; Pred. No. 5.4e-68;
Matches 451; Conservative 2; Mismatches 167; Indels 25; Gaps 4;

QY	16	CGTCCCTCCCTCTCCACTGCAGCTTCTTCTAGCCCGAACCACTTCTTCTCTGCT	75
DB	32	CCTGCTCTCTCTCTCCACAGCCCTTCTTCTTCTCAGCCACTTCTTCTTCTGCT	91
QY	76	TGTTTCTCTCTAGGCGGGAAGCTGAGTTCAGAGCCAGCCAGGCGGAGCAGCTC	135
DB	92	TCACCTCCAGTGCAGTGAAGAGGTAACCGGGTCCAGACCCAGCGGCGCGAGTCT	151

QY 136 TTCAGTGAAGAGCAACATCGAGGGTCAGCAATGAACGTGGAGCATGAGGTTAACTT 195
 Db 152 CCAGCGGGAAGAAACCGGAGAGGAGGAGCAATGAATGTGATCAGGAGTTAACTT 211
 QY 196 CCTGGTGGAGAAATTCATCGTCTGGGTTCACAAATGCGCATGGGAACAGTGTGAA 255
 Db 212 CTAGTGGAGAAATTCATCGTCTGGGTTCACAAATGCGCATGGGAACAGTGTGAA 271
 QY 256 GTTGGGGTCTCTTCGAGACGAGATGCGCAATCTCTTTGAAGCGTTGGTGGGAAC 315
 Db 272 ATTTGGGGTCTCTTCGAGATGAATAATGTGCAACCTCTTTGAAGCATTTGGTAGAAC 331
 QY 316 TCTGAAGCCGCAAAACGAAGGAATGTTACGTACGAGGAGAGCTGCTTTGCAAGG 375
 Db 332 TCTTAAAGCTGCAAAACGAAGGAATGTTAAACATATCCAGGAGAGCTGCTTTCGCAAGG 391
 QY 376 TGTTCATGATGATGTTGACATGTTGATGCTGCAAGATTAATGTGTTTGCAGATCTGGGG 435
 Db 392 TGTTCATGATGATGTTGACATGTTGATGCTGCAAGATTAATGTGTTTACATATCTTTAT 451
 QY 436 GTA-----TCTGGTAAACTGGAATTAATGAATTAAGTTAAAGGACAAACAT---G 477
 Db 452 GTACTGCCATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 510
 QY 478 AAGTTCCTTATGATGTTTATGACCTTTGATGACCTTTGATGACCTTTGATGACCTTTG 531
 Db 511 ACATACCTTAATGATGTTTATGACCTTTGATGACCTTTGATGACCTTTGATGACCTTTG 570
 QY 532 AGTCCCTGTTTATGACCTTTGATGACCTTTGATGACCTTTGATGACCTTTGATGACCTTT 591
 Db 571 CTGTCCTTTTATGACCTTTGATGACCTTTGATGACCTTTGATGACCTTTGATGACCTTT 630
 QY 592 TTTTCTTTTCTTAAAGAGTGAATCGGAGAGCTAGGCAATAAATA 636
 Db 631 ATTTTCTMAGNATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 675

RESULT 2

US-09-920-300A-939/c
 ; Sequence 939, Application US/09920300A
 ; Patent No. US20020136728A1
 ; GENERAL INFORMATION:
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Meagher, Madeleine Joy
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Secretist, Heather
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121.547
 ; CURRENT APPLICATION NUMBER: US/09/920,300A
 ; CURRENT FILING DATE: 2001-07-31
 ; NUMBER OF SEQ ID NOS: 1789
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 939
 ; LENGTH: 552
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-920-300A-939

Query Match 26.5%; Score 206.6; DB 10; Length 552;
 Best Local Similarity 68.7%; Pred. No. 4e-46;
 Matches 366; Conservative 0; Mismatches 144; Indels 23; Gaps 5;

QY 263 GTCTCTTCCAGACGACAGATGTCCTTTGAGCGTTGGTGGGAACTCTGAA 322
 Db 552 GTCTCTTCCGTGATGATAATGTCCTTTGAGCATTTGGTAGCAACTCTTAA 493
 QY 323 GCCCAAAACGAAGAGATGTTACGTACGAGGAGCTGCTTTTGAAGGTTGTTAT 382
 Db 492 GCTCAAAACGAAGAGATGTTACATATCCAGGAGCTGCTTCTGCAAGGTTGTTAT 433
 QY 383 GATGATGTTGACATGTTGCTGCAAGATTAATGTGTTGCAAGTCTGGGGTA---- 438

Db 432 GATGATGTTGACATTAATTAATCTGCAAGATTAATGTGGTTTACATATCTTTTATGTACTGC 373
 QY 439 -----TCTGTAACTGGATAATTAAGTTTAAAGGACAAACAT---GAAGTTTC 484
 Db 372 CATTTTCTGTTCTGTTAACTGGAT-ATAAGTGAAGAACAAACATTTGACATACT 314
 QY 485 TTATGATTTTTTATAGACCTTTGTAACAAAGGGGA--CTTGTGAGAGTCTCTGTTT 542
 Db 313 TAATGATTTTTTATAGAACTTTGTAACGAAAGGAGATTCATGTTTATAGAGCTGCTCT 254
 QY 543 TATACCTTGGGACAAACATTAATGTAATAAATAACAAACCTCTTATTTTTTTTTC 602
 Db 253 TTTTATATCTGTTGAAGAAATCTATGATGATGCTATATAAATAAATCTTATTTTTC 194
 QY 603 TTAGAAGTAACTCGGAGAGCTAGGCAATAAATGTTTTCAGAGGTGCGAAAGCTTT 662
 Db 193 TCAGGAATCTGTTAGGATTCGAGCATGAGATTTTTCGCGGCGAAGGATGGGANTG 134
 QY 663 TGTGTTTCTTAAACCATCTCTAGTCT--CTGCCACACTGCACACTCCGTCAAAGTGAAG 720
 Db 133 TTTGTTCAATAAATAATAGACATTTCTATAGATATTTGACATTTCTGCGAAAGCAACAG 74
 QY 721 CGAACTAAGACCAACTCGGTTGGGAAATATATGTTTATGTAATAAATAA 773
 Db 73 CAARCTGAAGACCAACTCTATGAGAAATATATGATGTTTATGTAATAAAGA 21

RESULT 3

US-10-033-528-939/c
 ; Sequence 939, Application US/10033528
 ; Patent No. US20020131971A1
 ; GENERAL INFORMATION:
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Meagher, Madeleine Joy
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Secretist, Heather
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121.547C1
 ; CURRENT APPLICATION NUMBER: US/10/033,528
 ; CURRENT FILING DATE: 2001-12-26
 ; NUMBER OF SEQ ID NOS: 1896
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 939
 ; LENGTH: 552
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-033-528-939

Query Match 26.5%; Score 205.6; DB 12; Length 552;
 Best Local Similarity 68.7%; Pred. No. 4e-46;
 Matches 366; Conservative 0; Mismatches 144; Indels 23; Gaps 5;

QY 263 GTCTCTTCCAGACGACAGATGTCCTTTGAGCGTTGGTGGGAACTCTGAA 322
 Db 552 GTCTCTTCCGTGATGATAATGTCCTTTGAGCATTTGGTAGCAACTCTTAA 493
 QY 323 GCCCAAAACGAAGAGATGTTACGTACGAGGAGCTGCTTTTGAAGGTTGTTAT 382
 Db 492 GCTCAAAACGAAGAGATGTTAATATCCAGGAGCTGCTTCTGCAAGGTTGTTAT 433
 QY 383 GATGATGTTGACATGTTGCTGCAAGATTAATGTGTTGCAAGTCTGGGGTA---- 438
 Db 432 GATGATGTTGACATTAATTAATGTTTAAAGGACAAACATTTGACATACT 373
 QY 439 -----TCTGTAACTGGATAATTAAGTTTAAAGGACAAACAT---GAAGTTTC 484
 Db 372 CATTTTCTGTTCTGTTAACTGGAT-ATAAGTGAAGAACAAACATTTGACATACT 314
 QY 485 TTATGATTTTTTATAGACCTTTGTAACAAAGGGGA--CTTGTGAGAGTCTCTGTTT 542
 Db 313 TAATGATTTTTTATAGAACTTTGTAACGAAAGGAGATTCATGTTTATAGAGCTGCTCT 254

Qy	543	TATACCTGGGAGCAAAACATTACAATGTAAAAATAAACAAAACCTGTATTTTTTTTTTTC	602
Db	253	TTTTTATATCTTGAATGAAAAATCTATGATGCTATAAAATAAATCCTATATATTTTC	194
Qy	603	TTAAGAAGGTAATCGGAGACCTAGGCAATAAAATGTTTTCAGAGGTGGCAAAAACCTTT	662
Db	193	TCAGGAATCTGGTTAGGAATGTCAGGCAATGAGATTTTTCGGGGCAAGGATGGGAATG	134
Qy	663	TGTTTTCTTAAACCATCTTCTAGTCT--CTGCCACACTTGACACTCGCTCAAACTGAGAAG	720
Db	133	TTTCTTCATAAAATAATTAGACATTTTCTATAGATATTTGCATCTTCGCGAAGCAACAAG	74
Qy	721	CGAACTAAGACCACTGCGGTGGAAAAATATATGTTTATGTATATAAAAAAA	773
Db	73	CAAACTGAAGACCAACTCCTATGAGAAATATATGATGTTTATGTATAAAGA	21

RESULT 4

US-09-925-299-318

; Sequence 318, Application US/09925299

; Patent No. US20020055627A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA102

; CURRENT APPLICATION NUMBER: US/09/925,299

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05883

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 1556

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 318

; LENGTH: 388

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc_feature

; LOCATION: (2)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc_feature

; LOCATION: (3)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc_feature

; LOCATION: (14)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc_feature

; LOCATION: (143)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc_feature

; LOCATION: (144)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc_feature

; LOCATION: (146)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc_feature

; LOCATION: (150)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc_feature

; LOCATION: (154)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc_feature

; LOCATION: (159)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc_feature

; LOCATION: (160)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc_feature

; LOCATION: (163)


```
QY 551 GGACGACAAACATTAACATGTAATAAATAAACAACCTCTTATTTTCTTAAAGAG 610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 233 TCTTGAAGAAATCTATGATGATGCTATATAAATAAATCTATATTTCTCAGGAAT 174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 611 GTAATCGGAGAGCTAGGCAATAAATGTTTTCAGAGGTGCGAAAGCTTTTGTCT 670
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 CTGCTTAGGAATTCGAGCAATGAGATTTTTCGGGCGAGGATGGAAATGTTGTCA 114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 671 TAAACCATTTCTAGTCT--CTGCCACACTTGACACTCCGTCAGAGTGAAGCGAACTAA 728
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 TAAATNATTAGACATTTTCTATAGATATTTGACATTTCTGCAAGCAACAAGCAAACTGA 54
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 729 AGACCAACTCGGTCGGAATAATATTATGTTATGTAATAAATAA 773
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 53 AGACCAACTCTATGAGAAATATTATGATGTTATGTAATAAAGA 9
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
US-09-880-107-687/c
; Sequence 687, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 687
; LENGTH: 406
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA287347
US-09-880-107-687

Query Match 11.8%; Score 92.2; DB 10; Length 406;
Best Local Similarity 62.0%; Pred. No. 1.9e-15;
Matches 214; Conservative 0; Mismatches 123; Indels 8; Gaps 4;

QY 436 GTATCTGTTAACTGGTAATTAAGTTAAGGACAAACAT---GAAGTTCCTTATGAT 492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 380 GTTCTGGTAACTGGAAAT-ATAAGTGAAGAAACAACATTTGAACATACCTTAATGAT 322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 493 TTTTATAGACCTTTGTAACAAAGGGGACT--TGTGAGAGTCCTGTTTATACCTT 550
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 321 TTTTATAGACTTGTAAACGAAGGATTCATGTTTGAAGTCTGCTCTTTTATA 262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 551 GGAGCAAAACATTACATGTAAATAAATAAACAACCTGTTATTTTCTTAAAGAG 610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 261 TCTTGAAGAAATCTATGATGATGCTATATAAATAAATCTTATTTCTCAGGAAT 202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 611 GTAATCGGAGAGCTAGGCAATAAATGTTTTCAGAGGTGCGAAAGCTTTTGTCT 670
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 CTGCTTAGGAATTCGAGCAATGAGATTTTTCGGGCGAGGATGGAAATGTTGTCA 142
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 671 TAAACCATTTCTAGTCT--CTGCCACACTTGACACTCCGTCAGAGTGAAGCGAACTAA 728
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 141 TAAATNATTAGACATTTCTATAGATATTTGACATTTCTGCAAGCAACAAGCAAACTGA 82
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 729 AGACCAACTCGGTCGGAATAATATTATGTTATGTAATAAATAA 773
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 81 AGACCAACTCTATGAGAAATATTATGATGTTATGTAATAAAGA 37
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 9
US-09-294-093B-2020
; Sequence 2020, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 2020
; LENGTH: 268
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700345395H1
US-09-294-093B-2020

Query Match 11.2%; Score 87.4; DB 10; Length 268;
Best Local Similarity 65.8%; Pred. No. 3.1e-14;
Matches 127; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 167 GCAATGAACGTGGAGCATGAGTAACTTCCTGCTGGTGGAGAAATTCATGCTGGGTCC 226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 GCCATGAACGTGGAGGAGGCTCGGGAGGCTCAAGGAGGAGATCCATAGGCTCGGCCAG 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 227 AAAATGCGGATGGGAACACTGAGTGAAGTTGGGTCTCTTCCAAAGACGACATGT 286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 CAGCAGCCCGATGGCTCTTACAAGGTCAAGTTTGGCGTTCTTCAACGACGATCGTGT 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 287 GCCAATCTCTTTGAACGTTGGTGGAACTCTGAAAGCCGCAAAACGAAGAGATTGTT 346
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 GCAATATCTTTGAAGCACTAGTTGGCACCTTGGACCTTGAAGGAGGAGGAGATCTTG 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 347 ACGTACGAGGAG 359
    ||| ||| ||| |||
Db 256 ACCTACGAAGGCG 268
    ||| ||| ||| |||

RESULT 10
US-09-294-093B-4150
; Sequence 4150, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 4150
; LENGTH: 289
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700353201H1
US-09-294-093B-4150
```



```

; PRIOR APPLICATION NUMBER: US 60/237,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO: 3235
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-3235

```

Query Match	4.9%	Score 38;	DB 9;	Length 2000;
Best Local Similarity	47.8%;	Pred. No. 1.6;		
Matches 110;	Conservative	0;	Mismatches 120;	Indels 0;
				Gaps 0;

QY	448	CTCGAATAATTAAGTTTAAAGGACAAACATGAAGTTCCTTATGATATTTTATAGACCTTTG	507
Db	25	CTTTAAAAATCTATGTGAAAATACCTTTTAAAAATCCCTGTAAATCTTATAAAACCCCTA	84
QY	508	TAAACAAAAGGGGACATGTGTGAGAACTCCTGTGTTTTATACCTTGGAGCAAAACATTCACAA	567
Db	85	TAAATCCCTTGTAAATCTTTTAAAACTCTTTTAAAAATCCCTGTGAATCTATTTCCTCCG	144
QY	568	TGTAAAAATAAACAAAACCTGTATTATTTTTTTTCTTAAGAAAGGTAATCGGAGACGTAG	627
Db	145	CAAAAAAATAAATTCGCGGGAATAATATGTTTGAACAACAAATTTGGCGGGAAATTTAT	204
QY	628	GCNATAAATGTTTTTCAGAGTTCGGAAGAGCTTTTGTCTTTTCTTAACAACA	677
Db	205	TTTGAATATTTTTTGGCGAAAAAATATTAATTTTTTCATTTGTGTAACTTA	254

RESULT 14
US-10-046-935-342/c
; Sequence 342, Application US/10046935
; Patent No. US20020156011A1

```

: GENERAL INFORMATION:
: APPLICANT: Jiang, Yuxiu
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Secrist, Heather
: APPLICANT: Wang, Aijun
: APPLICANT: Stolk, John A.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: OF INVENTION: 121 AND DIAGNOSIS OF COLON CANCER
: FILE REFERENCE: 210121.527C1
: CURRENT APPLICATION NUMBER: US/10/046,935
: CURRENT FILING DATE: 2002-01-15
: NUMBER OF SEQ ID NOS: 2239
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 342
: LENGTH: 338
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-046-935-342

```

Query Match	4.8%	Score 37.6;	DB 9;	Length 338;
Best Local Similarity	53.4%;	Pred. No. 0.85;		
Matches 79;	Conservative	0;	Mismatches	69;
			Indels	0;
			Gaps	0;

[illegible]

QY	609	AGGTAATCGGGAGACGTAGGCAATAAA	636
Db	205	AGGCTTCTCCAAATCGTGGGAATAACA	178

```

RESULT 15
US-09-878-178-342/c
: Sequence 342, Application US/09878178
: Patent No. US20020177552A1
: GENERAL INFORMATION:
: APPLICANT: Jiang, Yuqiu
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Sericist, Heather
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
: FILE REFERENCE: 210121.527

```

```

, FILE REFERENCE: 210121.527
, CURRENT APPLICATION NUMBER: US/09/878,178
, CURRENT FILING DATE: 2001-06-08
, NUMBER OF SEQ ID NOS: 2237
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO: 342
, LENGTH: 338
, TYPE: DNA
, ORGANISM: Homo sapien
US-09-878-178-342

```

Query Match 4.8%; Score 37.6; DB 9; Length 338;
Best Local Similarity 53.4%; Pred. No. 0.85;
Matches 79; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy	489	GTATTTTNTAGACCTTGTAAACAAAGGGGACTGTTGAGAAAGTCCCTGTTTTTATACC	548
Db	325	GTTATATATAATTAGTGTAGAAAAAATTTTATTTATTAACACAGTCTTTTTTGGTCC	266
Qy	549	TGGAGCAAAACATTAACAATGTAAAAATAAACAAAACCTCTATTTTTTTTTCTTTAAGA	608
Db	265	ATCAACACATATCTCTTCAGCAATAATAAACACAATCTTTTTTTTTTTTTTTTTGGGT	206

QY	609	AGGTAATCGGGAGACGTAGGCAATAAAA	636
Db	205	AGGCTTCTCCAAATCGTGGGAATAAACA	178

```
Search completed: January 17, 2003, 02:04:09
Job time : 112 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 21:41:43 ; Search time 71 Seconds
(without alignments)
3369.125 Million cell updates/sec

Title: US-09-648-310-1

Perfect score: 780

Sequence: 1 ggcacagctctctctgtcc.....gtaataaaaaaaatcatgt 780

Scoring table: IDENTITY_NDC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*

5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	157	20.1	177	2	US-09-032-684-20
2	44	5.6	7218	1	US-08-232-463-14
3	40	5.1	611	4	US-09-385-982-376
4	38	4.9	720	1	US-08-117-083-23
5	38	4.9	1500	1	US-08-117-083-67
6	37	4.7	1818	4	US-09-357-206A-6
7	37	4.7	5253	4	US-09-357-206A-16
8	37	4.7	5483	4	US-09-357-206A-17
9	37	4.7	5586	4	US-09-357-206A-19
10	37	4.7	5816	4	US-09-357-206A-21
11	37	4.7	6095	4	US-09-357-206A-18
12	37	4.7	6325	4	US-09-357-206A-20
13	37	4.7	6428	4	US-09-357-206A-22
14	37	4.7	7400	1	US-08-261-563A-1
15	37	4.7	7400	5	PCT-US95-07754A-1
16	37	4.7	10881	4	US-09-357-206A-9
17	37	4.7	12286	4	US-09-357-206A-1
18	36.8	4.7	3588	1	US-08-197-792-32
19	36.8	4.7	3588	1	US-08-459-850-32
20	36.8	4.7	3588	1	US-08-459-214-32
21	35.4	4.5	907	1	US-09-664-596B-23
22	35.4	4.5	907	2	US-08-739-775-1
23	35	4.5	4880	4	US-09-402-929-5
24	35	4.5	168575	4	US-09-426-290-1
25	34.4	4.4	4291	2	US-08-417-210A-81
26	34.4	4.4	5852	4	US-07-867-106-2
27	34.2	4.4	2981	1	US-08-257-073-2

28 34.2 4.4 2981 2 US-08-184-009-119 Sequence 119, App
29 34.2 4.4 2981 2 US-08-458-356-119 Sequence 119, App
30 34.2 4.4 2981 4 US-08-460-736-119 Sequence 119, App
31 34.2 4.4 3107 4 US-08-213-419B-1 Sequence 1, Appl
32 34.2 4.4 5506 4 US-09-004-838-93 Sequence 93, Appl
33 34.2 4.4 6124 4 US-08-213-419B-3 Sequence 3, Appl
34 34.2 4.4 7641 1 US-07-924-028A-5 Sequence 5, Appl
35 34.2 4.4 10607 1 US-08-078-090-3 Sequence 3, Appl
36 33.6 4.4 1713 4 US-09-134-001C-135 Sequence 135, App
37 33.6 4.3 7218 1 US-08-232-463-14 Sequence 14, Appl
38 33.4 4.3 1511 1 US-07-991-867B-8 Sequence 8, Appl
39 33.4 4.3 1511 1 US-08-107-755A-8 Sequence 8, Appl
40 33.4 4.3 1511 2 US-08-544-332-8 Sequence 8, Appl
41 33.4 4.3 1511 4 US-09-370-861A-8 Sequence 8, Appl
42 33.2 4.3 606 4 US-09-328-111-132 Sequence 132, App
43 33 4.2 1208 4 US-09-461-474-11 Sequence 11, Appl
44 33 4.2 2196 1 US-08-090-523-29 Sequence 29, Appl
45 33 4.2 2196 1 US-08-334-639-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-032-684-20
; Sequence 20, Application US/09032684
; Patent No. 5882874
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; TITLE OF INVENTION: RECIPROCAL SUBTRACTION DIFFERENTIAL
; TITLE OF INVENTION: DISPLAY
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,684
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 55551/JPW/AMG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-032-684-20

Query Match 20.1%; Score 157; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 9e-37;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 624 GTAGGCATATAATGTTTCAGAGGTGCGAAAGCTTTTGTCTTAAACCACTCTTA 683
|||||

DB 1 GTAGGCATATAATGTTTCAGAGGTGCGAAAGCTTTTGTCTTAAACCACTCTTA 60
|||||

QY 684 GTCTCTGCCACACTTGACACTCCGTCAAAGTGAGAGCGCACTAAAGACCACTCGGTG 743
|||||

```
Db 61 GTCTCCACACTTGACACTCGCTAAAGTGAGAGGAACTAAAGACCAACTGCGGTG 120
QY 744 GAAATATTATGTTATGTAATAAAAAAATCATGT 780
Db 121 GAAATATTATGTTATGTAATAAAAAAATCATGT 157

RESULT 2
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFELINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match 5.6%; Score 44; DB 1; Length 7218;
Best Local Similarity 14.1%; Pred. No. 0.0066;
Matches 80; Conservative 214; Mismatches 274; Indels 0; Gaps 0;

QY 78 TTCCTCCCTAGGCGGGAAGCTGAGTGCAGGGTTTCAGACCCAGCGGCGAGCAGCTCTT 137
Db 1441 TTGCTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1382
QY 138 CAGTGAGAGGAGGAGCAATCGGAGGTGACCAATGACGTGGAGCATGAGGTAACTCC 197
Db 1381 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1322
QY 198 TGGTGAGGAAATTCATCGTCTGGGTTCCAAAATCCGATGGAACTGAGTGTAGT 257
Db 1321 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1262
```

```
QY 258 TTGGGGTCTCTTCCAGACGACAGATGTGCCAATCTCTTTGAAGGTTGTTGGGAATC 317
Db 1261 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1202
QY 318 TCAAAGCCCAACCAAGCAAGATTTGTTACGTACGAGGAGAGCTGCTTTTCCAAAGTG 377
Db 1201 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1142
QY 378 TTCATGATGATGTGACATTTGATCTGCTCAAGATTAAATGCTGTTGTCAGATCTGGGGT 437
Db 1141 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1082
QY 438 ATCTGTAACTGGAATAATTAAGTTAAAGGACAAACATGAAGTCTCTTATGTTATTTA 497
Db 1081 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1022
QY 498 TAGACCTTTGTAACAAAGGAGGACTTGTGAGAACTCTCTGTTTATACCTTTGGAGCA 557
Db 1021 TGAGCGTATGCAACGAAGAGAAATAGTTATAGTACGCGCACTCGATGGGACATTTCA 962
QY 558 AACATTACAATGTAATAATAACAAACCTGTTATTTTTTTTCTTAAGAGGTAACTG 617
Db 961 ACGTAAACCGTTTAAATAATTTTGAATCTTATTCATTATCTGAAATGTTGTAATACT 902
QY 618 GGAGACGTAGGCAATAAATAATTTTTCAG 645
Db 901 AACTGCTGTGTATGCAATGCTTTAAG 874

RESULT 3
US-09-385-982-376
; Sequence 376, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 376
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(611)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-376

Query Match 5.1%; Score 40; DB 4; Length 611;
Best Local Similarity 62.6%; Pred. No. 0.029;
Matches 107; Conservative 0; Mismatches 54; Indels 10; Gaps 3;

QY 436 GTATCTGTAACACTGGAATAATTAAGTTAAAGGACAAACAT---GAAGTCTCTATGAT 492
Db 325 GTTTNGGTAACNGCAAT-ATAANGNGAAAGAACACNTTTGGAACATCTAATGAT 383
QY 493 TTTTATAGACCTTTTAAACAAAGGAGCTGT-----TGAGAGTCTCTCTTTTATA 546
Db 384 TTTTATAGACCTTTGNAACCAAGAGGATTCATGTTTANAAGTCTGCGCTTTTATA 443
QY 547 CTTTGGAGCAAAACATTAACAATGTAATAAATAACAAACCTGTTATTTT 597
Db 444 TCTTGGAGAAATATCTGTTGAGGCTTNTAATAATATCCCATTTTCT 494
```

```

RESULT 4
US-08-117-083-23
; Sequence 23, Application US/08117083
; Patent No. 5719054
; GENERAL INFORMATION:
; APPLICANT: Bournell, Michael E.
; APPLICANT: Ingilis, Stephen C.
; APPLICANT: Munro, Alan J.
; TITLE OF INVENTION: Recombinant Virus Vectors Encoding Human
; TITLE OF INVENTION: Papilloma Virus Proteins
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Walter H. Dreger
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/117,083
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Walter H.
; REGISTRATION NUMBER: 24,190
; REFERENCE/DOCKET NUMBER: A-58783
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 720 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-117-083-23

```

Query Match	4.9%	Score 38;	DB 1;	Length 720;
Best Local Similarity	49.5%;	Pred. No. 0.12;		
Matches	98;	Conservative 0;	Mismatches 100;	Indels 0; Gaps 0;
QY 351	ACGCAGGAGAGCTGCTTTTGC	AAAGTGTTCATGATGATGTTGACATTTGCTGTCGAAG	410	
Db 296	ATGCTGTTGTGCTGTGTTGC	AGATTTGGCCAGAAATCATGGTATATGGATACATAATGGTAT	355	
QY 411	ATTAATCTGGTTTCGACATCT	GGGGTATCTGGTAAACTTGGAAATAATAAGTTAAAGGAC	470	
Db 356	CAAGAAGTATTCTCCAGAT	GAATGGGTGCACATATAAAATTTTAATTAATGTAATAGAG	415	
QY 471	AAACATCAAGTTCCTTATG	TATTTTATAGACCTTTGTAACAAAGGGGACATTTGTTGAG	530	
Db 416	AACAAATAAATGAGTTG	TATATCATATAGACAATAACTAAACAATTAATAGTAACCTGT	475	
QY 531	AAGTCCTGTTTTTATACC	548		
Db 476	ATCTCTTTTTTAACTAAC	493		

RESULT 5
US-08-117-083-67
; Sequence 67, Application US/08117083
; Patent No. 571954
; GENERAL INFORMATION:
; APPLICANT: Boursnell, Michael E.

APPLICANT: Inglis, Stephen C.
 APPLICANT: Munro, Alan J.
 TITLE OF INVENTION: Recombinant Virus Vectors Encoding Human
 TITLE OF INVENTION: Papilloma Virus Proteins
 NUMBER OF SEQUENCES: 70
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Walter H. Dreger
 STREET: 4 Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: CA
 COUNTRY: USA
 ZIP: 94111
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/117,083
 FILING DATE: 10-SEP-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Dreger, Walter H.
 REGISTRATION NUMBER: 24,190
 REFERENCE/DOCKET NUMBER: A-58783
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-781-1989
 TELEFAX: 415-398-3249
 TELEX: 910 277299
 INFORMATION FOR SEQ ID NO: 67:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1500 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-117-083-67

Query Match 4.9%; Score 38; DB 1; Length 1500;
Best Local Similarity 49.5%; Pred. No. 0.18;
Matches 98; Conservative 0; Mismatches 100; Indels

[illegible]

```

RESULT 6
US-09-357-206A-6/C
: Sequence 6, Application US/09357206A
: Patent No. 637962
: GENERAL INFORMATION:
: APPLICANT: Dinesh-Kumar, S.
: APPLICANT: Baker, Barbara
: TITLE OF INVENTION: Pathogen Resistance
: FILE REFERENCE: 042250/191805 (3830-5)
: CURRENT APPLICATION NUMBER: US/09/357,206A
: CURRENT FILING DAPE: 1999-07-20
: PRIOR APPLICATION NUMBER: US 60/093,494
: PRIOR FILING DATE: 1998-07-20
: NUMBER OF SEQ ID NOS: 22

```

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 6

; LENGTH: 1818

; TYPE: DNA

; ORGANISM: Nicotiana glauca

US-09-357-206A-6

Query Match 4.7%; Score 37; DB 4; Length 1818;

Best Local Similarity 50.7%; Pred. No. 0.38;

Matches 115; Conservative 0; Mismatches 110; Indels 2; Gaps 1;

QY 390 TTGACATTGTTGTCGCAAGATTAAATGTTGCGAGATCTGGGGTATCTGGTAACT 449

Db 1490 TTGAATACAGAGGCCAACACATATTTTGTGTAATAGAAATTTATTTATATA 1431

QY 450 GGAATATTA--AGTTAAAGGACAAACATGAAGTCTCTATGTTATTTTATAGACCTTTG 507

Db 1430 TGGAAATTTACTTGTAAAGAACTGAATTAAGGTAAATTTATTTATATATATA 1371

QY 508 TAAACAAAGGGGACTGTTGAGAAGTCTGTTTATACCTTGGAGCAAAACATTACAA 567

Db 1370 AATTACAAAGGCTCTTAATGTCGTCGTTTATACCTTTTAAATAAATTTTACAC 1311

QY 568 TGTAAATAAATAACAAACCTGTTATTTTCTTTTCTTAAGAAGTAA 614

Db 1310 TAGAGAAATATATTTTAAAGTAAATTTCTTAATATTTAGGATTTA 1264

RESULT 7

US-09-357-206A-16/c

; Sequence 16, Application US/09357206A

; Patent No. 6372962

; GENERAL INFORMATION:

; APPLICANT: Dinesh-Kumar, S.

; APPLICANT: Baker, Barbara

; TITLE OF INVENTION: Pathogen Resistance in Plants using cDNA-N/Intron Constructs

; FILE REFERENCE: 042250/191805 (5830-5)

; CURRENT APPLICATION NUMBER: US/09/357,206A

; CURRENT FILING DATE: 1999-07-20

; PRIOR APPLICATION NUMBER: US 60/093,494

; PRIOR FILING DATE: 1998-07-20

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 16

; LENGTH: 5253

; TYPE: DNA

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: cDNA-N/intron construct: E1-E2-E3-I3-E4-E5

US-09-357-206A-16

Query Match 4.7%; Score 37; DB 4; Length 5253;

Best Local Similarity 50.7%; Pred. No. 0.64;

Matches 115; Conservative 0; Mismatches 110; Indels 2; Gaps 1;

QY 390 TTGACATTGTTGTCGCAAGATTAAATGTTGCGAGATCTGGGGTATCTGGTAACT 449

Db 3338 TTGAATACAGAGGCCAACACATATTTTGTGTAATAGAAATTTATTTTATATA 3279

QY 450 GGAATATTA--AGTTAAAGGACAAACATGAAGTCTCTATGTTATTTTATAGACCTTTG 507

Db 3278 TGGAAATTTACTTGTAAAGAACTGAATTAAGGTAAATTTATTTATATATATA 3219

QY 508 TAAACAAAGGGGACTGTTGAGAAGTCTCTTTTATACCTTGGAGCAAAACATTACAA 567

Db 3218 AATTACAAAGGCTCTTAATGTCGTCGTTTATACCTTTTAAATAAATTTTACAC 3159

QY 568 TGTAAATAAATAACAAACCTGTTATTTTCTTTTCTTAAGAAGTAA 614

Db 3158 TAGACGAAATATATTTTAAAGTAAATTTTCTTAATATTTAGGATTTA 3112

RESULT 8

US-09-357-206A-17/c

; Sequence 17, Application US/09357206A

; Patent No. 6372962

; GENERAL INFORMATION:

; APPLICANT: Dinesh-Kumar, S.

; APPLICANT: Baker, Barbara

; TITLE OF INVENTION: Pathogen Resistance in Plants using cDNA-N/Intron Constructs

; FILE REFERENCE: 042250/191805 (5830-5)

; CURRENT APPLICATION NUMBER: US/09/357,206A

; CURRENT FILING DATE: 1999-07-20

; PRIOR APPLICATION NUMBER: US 60/093,494

; PRIOR FILING DATE: 1998-07-20

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 17

; LENGTH: 5483

; TYPE: DNA

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: cDNA-N/intron construct: E1-I1-E2-E3-I3-E4-E5

US-09-357-206A-17

Query Match 4.7%; Score 37; DB 4; Length 5483;

Best Local Similarity 50.7%; Pred. No. 0.65;

Matches 115; Conservative 0; Mismatches 110; Indels 2; Gaps 1;

QY 390 TTGACATTGTTGTCGCAAGATTAAATGTTGCGAGATCTGGGGTATCTGGTAACT 449

Db 3568 TTGAATACAGAGGCCAACACATATTTTGTGTAATAGAAATTTATTTATATA 3509

QY 450 GGAATATTA--AGTTAAAGGACAAACATGAAGTCTCTTATGTTATTTTATAGACCTTTG 507

Db 3508 TGGAAATTTACTTGTAAAGAACTGAATTAAGGTAAATTTATTTATATATATA 3449

QY 508 TAAACAAAGGGGACTGTTGAGAAGTCTCTGTTTATACCTTGGAGCAAAACATTACAA 567

Db 3448 AATTACAAAGGCTCTTAATGTCGTCGTTTATACCTTTTAAATAAATTTTACAC 3389

QY 568 TGTAAATAAATAACAAACCTGTTATTTTCTTTTCTTAAGAAGTAA 614

Db 3388 TAGACGAAATATATTTTAAAGTAAATTTCTAATATTTAGGATTTA 3342

RESULT 9

US-09-357-206A-19/c

; Sequence 19, Application US/09357206A

; Patent No. 6372962

; GENERAL INFORMATION:

; APPLICANT: Dinesh-Kumar, S.

; APPLICANT: Baker, Barbara

; TITLE OF INVENTION: Pathogen Resistance in Plants using cDNA-N/Intron Constructs

; FILE REFERENCE: 042250/191805 (5830-5)

; CURRENT APPLICATION NUMBER: US/09/357,206A

; CURRENT FILING DATE: 1999-07-20

; PRIOR APPLICATION NUMBER: US 60/093,494

; PRIOR FILING DATE: 1998-07-20

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 19

; LENGTH: 5586

; TYPE: DNA

; ORGANISM: artificial

; FEATURE:

; OTHER INFORMATION: cDNA-N/intron construct: E1-E2-E3-I3-E4-I4-E5

US-09-357-206A-19

Query Match 4.7%; Score 37; DB 4; Length 5586;

Best Local Similarity 50.7%; Pred. No. 0.65;

Matches 115; Conservative 0; Mismatches 110; Indels 2; Gaps 1;

QY 390 TTGACATTGTTGTCGCAAGATTAAATGTTGCGAGATCTGGGGTATCTGGTAACT 449

Db 3338 TTGAATACAGAGGCCAACACATATTTTGTGTAATAGAAATTTATTTTATATA 3279

[illegible]

```

RESULT 10
US-09-357-206A-21/c
; Sequence 21, Application US/09357206A
; Patent No. 6372962
; GENERAL INFORMATION:
; APPLICANT: Dinesh-Kumar, S.
; APPLICANT: Baker, Barbara
; TITLE OF INVENTION: Pathogen Resistance in Plants using cDNA-N/Intron Constructs
; FILE REFERENCE: 042250/191805 (5830-5)
; CURRENT APPLICATION NUMBER: US/09/357,206A
; CURRENT FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: US 60/093,494
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 5816
; TYPE: DNA
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: cDNA-N/Intron construct: EI-11-E2-E3-I3-E4-I4-E5
US-09-357-206A-21

```

```

RESULT 11
US-09-357-206A-18/c
; Sequence 18, Application US/09357206A
; Patent No. 6372962
; GENERAL INFORMATION:
; APPLICANT: Dinesh-Kumar, S.
; APPLICANT: Baker, Barbara
; TITLE OF INVENTION: Pathogen Resistance in Plants using CDNA-N/Intron Constructs
; FILE REFERENCE: 042250/191805 (5830-5)
; CURRENT APPLICATION NUMBER: US/09/357,206A
; CURRENT FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: US 60/093,494
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0

```

```

; SEQ ID NO 18
;
; LENGTH: 6095
;
; TYPE: DNA
;
; ORGANISM: artificial
;
; FEATURE:
;
;   OTHER INFORMATION: cdna-N/intron construct: E1-E2-I2-E3-I3-E4-E5
US-09-357-206A-18

```

```

RESULT 12
US-09-357-206A-20/c
; Sequence 20, Application US/09357206A
; Patent No. 6372962
; GENERAL INFORMATION:
; APPLICANT: Dinesh-Kumar, S.
; APPLICANT: Baker, Barbara
; TITLE OF INVENTION: Pathogen Resistance in Plants using cDNA-N/Intron Constructs
; FILE REFERENCE: 042250/191805 (5830-5)
; CURRENT APPLICATION NUMBER: US/09/357,206A
; CURRENT FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: US 60/093,494
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 6325
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: cDNA-N/Intron construct: E1-I1-E2-I2-E3-I3-E4-E5
US-09-357-206A-20

```

RESULT 13

US-09-357-206A-22/c

Sequence 22, Application US/09357206A

Patent No. 6372962

GENERAL INFORMATION:

APPLICANT: Dinesh-Kumar, S.

APPLICANT: Baker, Barbara

TITLE OF INVENTION: Pathogen Resistance in Plants using cDNA-N/Intron Constructs

FILE REFERENCE: 042250/191805 (5830-5)

CURRENT APPLICATION NUMBER: US/09/357,206A

CURRENT FILING DATE: 1999-07-20

PRIOR APPLICATION NUMBER: US 60/093,494

PRIOR FILING DATE: 1998-07-20

NUMBER OF SEQ ID NOS: 22

SOFTWARE: PatentIn version 3.0

SEQ ID NO 22

LENGTH: 6428

TYPE: DNA

ORGANISM: artificial

FEATURE:

OTHER INFORMATION: cDNA-N/intron construct: E1-E2-I2-E3-I3-E4-I4-E5

US-09-357-206A-22

Query Match 4.7%; Score 37; DB 4; Length 6428;

Best Local Similarity 50.7%; Pred. No. 0.7;

Matches 115; Conservative 0; Mismatches 110; Indels 2; Gaps 1;

QY 390 TTGCACATTGTTCGCAAGATTAAATGCTGTTGCAGATCTGGGGTATCTCGTAAACT 449

DB 4180 TTGATTTACGAGGCCACACATATTTTGTGTGAATATAGAAATTTATTTTTATATA 4121

QY 450 GGAATAATTA--AGTTAAGGACCAACATGAAGTTCCTTATGTTATTTTATAGACCTTG 507

DB 4120 TGGAAATTTACTTGTTAAGAACTGAATAAAGGTAATTTATTTATATATATTA 4061

QY 508 TAAACAAAGGGACATGTTGAGAAGTCCTGTTTTTATACCTTGGAGCAAAACATTACAA 567

DB 4060 AAATTACAAAGGCTCTTAATGTCTGCTCTTTTACCCTTTTAAANAATAATTTTACAC 4001

QY 568 TGTAAAAATAACAAACCTGTTATTTTCTTTTCTTAAAGAGTAA 614

DB 4000 TAGACGAATATATTAGTTAATTTCTAATATTAGGATTTTA 3954

RESULT 14

US-08-261-663A-1/c

Sequence 1, Application US/08261663A

Patent No. 5571706

GENERAL INFORMATION:

APPLICANT: Baker, Barbara J

APPLICANT: Whitham, Steven A

TITLE OF INVENTION: Plant Virus Resistance Gene and Methods

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Margaret A. Connor, USDA-ARS

STREET: 800 Buchanan Street

CITY: Albany

STATE: CA

COUNTRY: USA

ZIP: 94710

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/261,663A

FILING DATE:

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Connor, Margaret A

REGISTRATION NUMBER: 30043

REFERENCE/DOCKET NUMBER: 0094.94

```
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/07754A
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Condor, Margaret A
;; REGISTRATION NUMBER: 30043
;; REFERENCE/DOCKET NUMBER: 0094.94
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (510) 559-6067
;; TELEFAX: (510) 559-5777
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 7400 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; ORIGINAL SOURCE:
;; ORGANISM: Nicotiana glauca
;; TISSUE TYPE: leaf
;; FEATURE:
;; NAME/KEY: exon
;; LOCATION: join(294..772, 1003..2098, 2941..3213, 5032..6600,
;; LOCATION: 6934..6951)
;; FEATURE:
;; NAME/KEY: intron
;; LOCATION: 773..1002
;; FEATURE:
;; NAME/KEY: intron
;; LOCATION: 2099..2940
;; FEATURE:
;; NAME/KEY: intron
;; LOCATION: 3214..5031
;; FEATURE:
;; NAME/KEY: intron
;; LOCATION: 6601..6933
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: join(294..772, 1003..2098, 2941..3213, 5032..6600,
;; LOCATION: 6934..6951)
;; PCT-US95-07754A-1
```

```
Query Match 4.7%; Score 37; DB 5; Length 7400;
Best Local Similarity 50.7%; Pred. No. 0.75;
Matches 115; Conservative 0; Mismatches 110; Indels 2; Gaps 1;

Qy 390 TTGACATTGTTGCTGCAAGATTAAATGTTGTCAGATCTGGGGTATCTGGTAAACT 449
Db 4703 TTGAATTACGAGGGCCACACATTATTTTCTGTAATAGAAATTTATTTTATATA 4644

Qy 450 GGAATAAATA--AGTAAAGGACAAACATGAAGTCCCTATGTATTTTATAGACCTTTG 507
Db 4643 TCGAAATTTACTTGTAAAGAACTCAATAAGGTAATTTATATTTTATATATA 4584

Qy 508 TAAACAAAAGGGGACTTGTGAGAGTCTCTGTTTATACCTTGGAGCAAAACATTACAA 567
Db 4583 AAATTACAAAAGGCTCTTAATGCTGCTCTTTTACCCCTTTAAATAAATTTTACAC 4524

Qy 568 TGTAATAATAACAAACCTGTTATTTTCTTAAAGAGGTAA 614
Db 4523 TAGACCAATTAATTTTAAAGTTAATTTCTTAATATTTAGGATTTTA 4477
```

Search completed: January 17, 2003, 02:02:32
Job time : 100 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 21:51:23 ; Search time 143 Seconds
(without alignments)
75.478 Million cell updates/sec

Title: US-09-648-310-2
Perfect score: 406
Sequence: 1 MNVEHEVLLVEIHLGSK.....AGELLQGVHDDVDIVLLQD 81

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	406	100.0	81	23	AAU76532
2	390	96.1	81	23	AAU76533
3	390	96.1	142	21	AAB36907
4	375	92.4	83	21	AAB32056
5	302	74.4	87	21	AAG58854
6	286	70.4	83	21	AAB32055
7	286	70.4	95	21	AAG58354
8	286	70.4	161	21	AAG58353
9	232	57.1	92	20	AAV39325
10	158	38.9	77	22	ABBI4740

Rat Progression Su
Human Progression
Human prostate can
Human secreted pro
Zea mays protein f
Human secreted pro
Arabidopsis thalia
Arabidopsis thalia
PSgen13 protein.
Human nervous syst

11	106	26.1	456	22	ABG27995	Novel human diagno
12	81.5	20.1	183	22	ABB58876	Drosophila melanog
13	73	18.0	399	22	ABB60049	Drosophila melanog
14	66.5	16.4	116	22	AAB46383	H. pylori HPC024 p
15	65	16.0	324	21	AAV91948	Human cytoskeleton
16	65	16.0	407	23	ABB97460	Novel human protei
17	64	15.8	671	23	ABP51430	Human MDDP SEQ ID
18	63.5	15.6	496	23	ABB92016	Herbicidally activ
19	63	15.5	575	22	AAU50377	Propionibacterium
20	62.5	15.4	491	23	ABB92013	Herbicidally activ
21	62.5	15.4	496	23	ABB92014	Herbicidally activ
22	62.5	15.4	565	19	AAW98798	H. pylori GHPO 128
23	62.5	15.4	565	19	AAW71554	Helicobacter polyp
24	62.5	15.4	565	22	AAW46327	H. pylori HPS024 p
25	62	15.3	137	22	AAO06434	Human polypeptide
26	61.5	15.1	676	23	ABP41814	Human ovarian anti
27	61	15.0	556	20	RAY38731	Neisseria meningit
28	60.5	14.9	101	21	AAV93885	Amino acid sequenc
29	60.5	14.9	387	21	AAG21474	Arabidopsis thalia
30	60.5	14.9	387	21	AAG32491	Arabidopsis thalia
31	60.5	14.9	387	21	AAG49686	Arabidopsis thalia
32	60.5	14.9	398	21	AAG21473	Arabidopsis thalia
33	60.5	14.9	398	21	AAG32490	Arabidopsis thalia
34	60.5	14.9	398	21	AAG49685	Arabidopsis thalia
35	60.5	14.9	573	17	AAW00164	Myosin heavy chain
36	60.5	14.9	573	21	AAG21472	Arabidopsis thalia
37	60.5	14.9	573	21	AAG32489	Arabidopsis thalia
38	60.5	14.9	573	21	AAG49684	Arabidopsis thalia
39	60.5	14.9	798	22	ABB64631	Drosophila melanog
40	60.5	14.9	1113	22	ABB64731	Drosophila melanog
41	60	14.8	1305	21	AAB12875	Murine JNK3 bindin
42	60	14.8	1314	21	AAB12876	Murine JNK3 bindin
43	60	14.8	1336	21	AAB12878	Murine JNK3 bindin
44	60	14.8	1337	21	AAB12877	Murine JNK3 bindin
45	59.5	14.7	157	22	AAO01991	Human polypeptide

ALIGNMENTS

RESULT 1
AAU76532
ID AAU76532 standard; Protein; 81 AA.
XX
AC AAU76532;
XX
DF 05-JUN-2002 (first entry)
XX
DE Rat Progression Suppressed Gene 13 (rPSGen 13).
XX
KW Rat; Progression Suppressed Gene 13; rPSGen 13; cytostatic; cancer;
KW blood vessel; nasopharyngeal tumour; thyroid tumour; leukaemia;
KW lymphoma; breast; lung; prostate; ovary; colon.
XX
OS Rattus sp.
XX
PN WO200216419-A2.
XX
PD 28-FEB-2002.
XX
PF 27-AUG-2001; 2001WO-US26795.
XX
PR 25-AUG-2000; 2000US-0648310.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Fisher PB, Kang D, Su Z;
XX
DR WPI; 2002-280914/32.
DR N-PSDB; ABK11085.
XX
PT New rat and human Progression Suppressed Gene 13 for preventing the growth of cancer cells and/or new blood vessels, and for treating

PT patients suffering from a cancer -
 XX Claim 28; Fig 1; 53pp; English.
 XX
 CC The invention relates to novel isolated nucleic acids which encode a
 CC rat or human Progression Suppressed Gene 13 (PSGen 13) protein. The
 CC nucleic acids are useful for preventing the growth of cancer cells and/or
 CC new blood vessels, and for treating patients suffering from a cancer,
 CC e.g. nasopharyngeal tumour, thyroid tumour, leukaemia, lymphoma, or
 CC cancer of the breast, lung, prostate, ovary or colon. PSGen 13 may also
 CC be used to suppress the transformed phenotype of a malignant cell.
 CC Administration of PSGen 13 gene or protein may result in a decrease in
 CC tumour mass, number of cancer cells, serum tumour marker, tumour
 CC metastasis, vascularisation, perfusion, or rate of tumour growth.
 CC Improved clinical symptoms, and/or increased patient survival. The
 CC present sequence represents the amino acid sequence of rat Progression
 CC Suppressed Gene 13 (rPSGen 13).
 XX
 SQ Sequence 81 AA;
 Query Match 100.0%; Score 406; DB 23; Length 81;
 Best Local Similarity 100.0%; Pred. No. 5.9e-46;
 Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MNVEHEVNLVEEIHRLGSKNADGKLSVKGVLFGVLFQDDRCANLFEALVGTLKAARRKIYV 60
 DB 1 MNVEHEVNLVEEIHRLGSKNADGKLSVKGVLFGVLFQDDRCANLFEALVGTLKAARRKIYV 60
 QY 61 YAGELLQGVHDDVDIVLLQD 81
 DB 61 YAGELLQGVHDDVDIVLLQD 81
 RESULT 2
 AAU76533
 ID AAU76533 standard; Protein; 81 AA.
 AC AAU76533;
 XX
 DT 05-JUN-2002 (first entry)
 XX
 DE Human Progression Suppressed Gene 13 (rPSGen 13).
 XX
 KW Human; Progression Suppressed Gene 13; rPSGen 13; cytostatic; cancer;
 KW Blood vessel; nasopharyngeal tumour; thyroid tumour; leukaemia;
 KW lymphoma; breast; lung; prostate; ovary; colon.
 XX
 OS Homo sapiens.
 XX
 PN WO200216419-A2.
 XX
 PD 28-FEB-2002.
 XX
 PF 27-AUG-2001; 2001WO-US26795.
 XX
 PR 25-AUG-2000; 2000US-0648310.
 XX
 PA (UYCO) UNIV COLUMBIA NEW YORK.
 XX
 PI Fisher PB, Kang D, Su Z;
 XX
 DR WPI; 2002-280914/32.
 DR N-PSDB; ABK11086.
 XX
 PT New rat and human Progression Suppressed Gene 13 for preventing the
 PT growth of cancer cells and/or new blood vessels, and for treating
 PT patients suffering from a cancer -
 XX
 PS Claim 30; Fig 2; 53pp; English.
 CC
 CC The invention relates to novel isolated nucleic acids which encode a
 CC rat or human Progression Suppressed Gene 13 (PSGen 13) protein. The
 CC nucleic acids are useful for preventing the growth of cancer cells and/or

CC new blood vessels, and for treating patients suffering from a cancer,
 CC e.g. nasopharyngeal tumour, thyroid tumour, leukaemia, lymphoma, or
 CC cancer of the breast, lung, prostate, ovary or colon. PSGen 13 may also
 CC be used to suppress the transformed phenotype of a malignant cell.
 CC Administration of PSGen 13 gene or protein may result in a decrease in
 CC tumour mass, number of cancer cells, serum tumour marker, tumour
 CC metastasis, vascularisation, perfusion, or rate of tumour growth.
 CC Improved clinical symptoms, and/or increased patient survival. The
 CC present sequence represents the amino acid sequence of human Progression
 CC Suppressed Gene 13 (hPSGen 13).
 XX
 SQ Sequence 81 AA;
 Query Match 96.1%; Score 390; DB 23; Length 81;
 Best Local Similarity 93.8%; Pred. No. 7.7e-44;
 Matches 76; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MNVEHEVNLVEEIHRLGSKNADGKLSVKGVLFGVLFQDDRCANLFEALVGTLKAARRKIYV 60
 DB 1 MNVEHEVNLVEEIHRLGSKNADGKLSVKGVLFGVLFQDDRCANLFEALVGTLKAARRKIYV 60
 QY 61 YAGELLQGVHDDVDIVLLQD 81
 DB 61 YAGELLQGVHDDVDIVLLQD 81
 RESULT 3
 AAB56907
 ID AAB56907 standard; Protein; 142 AA.
 AC AAB56907;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE Human prostate cancer antigen protein sequence SEQ ID NO:1485.
 XX
 KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
 KW vulnary; gastrointestinal; nephrotropic; antineoplastic; gynaecological;
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KW wound; infectious disease.
 XX
 OS Homo sapiens.
 XX
 PN WO2000055174-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05988.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2000-587513/55.
 DR N-PSDB; AAF16110.
 XX
 PT Prostate cancer associated gene sequences, referred to as prostate
 PT cancer antigens, useful for treatment, prevention, and diagnosis of
 PT disorders such as prostate cancer -
 XX
 PS Claim 11; Page 1922; 2338pp; English.
 CC
 CC AAF15566 to AAF16505 encode the human prostate cancer associated
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
 CC The prostate cancer antigens can have neuroprotective, cytostatic,
 CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
 CC nephrotropic, antineoplastic, gynaecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen

CC polynucleotides may be used for detection of prostate cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
 CC AAB57303 represent sequences used in the exemplification of the present
 CC invention.

XX Sequence 142 AA;

Query Match 96.1%; Score 390; DB 21; Length 142;
 Best Local Similarity 93.8%; Pred. No. 1.6e-43;
 Matches 76; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNVEHEVNLVVEIHRIGSKNADGKLSVKFGVLFQDDRCANLFEALVGTLLKAARRKIYV 60
 DB 62 MNVDHEVNLVVEIHRIGSKNADGKLSVKFGVLFQDDRCANLFEALVGTLLKAARRKIYV 121
 QY 61 YAGELLQGVHDDVDIVLQD 81
 DB 122 YPGELLQGVHDDVDIVLQD 142

RESULT 4

AAB32056
 ID AAB32056 standard; Protein; 83 AA.

XX AAB32056;

XX 14-FEB-2001 (first entry)

XX Human secreted protein BLAST search protein SEQ ID NO: 114.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.

OS Homo sapiens.

XX WO200058350-A1.

XX 05-OCT-2000.

XX 22-MAR-2000; 2000WO-US07483.

XX 26-MAR-1999; 99US-0126596.

XX 22-DEC-1999; 99US-0171552.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-602357/57.

XX Nucleic acid molecules encoding human secreted proteins, used in
 PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
 PT Parkinson's diseases and cancers -

PS Disclosure; Page 387; 423pp; English.

XX The invention relates to the isolation of genes AAC66410-C66458 encoding
 CC the human secreted proteins AAB32002-B32050. This sequence represents a
 CC fragment of the protein encoded by the gene given in the descriptor
 CC line. The sequence is used as a query sequence for doing BLASTX searches
 CC to determine homologous sequence to the protein. The genes and proteins
 CC are useful for preventing, ameliorating or treating medical conditions,
 CC e.g. by protein or gene therapy. The genes are isolated from a range of
 CC human tissues disclosed in the specification. The nucleic acids,
 CC proteins, antibodies and (ant)agonists are useful in the diagnosis,
 CC treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer,

CC and other cancers of the adrenal gland, bone, bone marrow, breast,
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders
 CC e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
 CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
 CC sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and
 CC (f) infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections.

XX Sequence 83 AA;

Query Match 92.4%; Score 375; DB 21; Length 83;
 Best Local Similarity 93.6%; Pred. No. 7.7e-42;
 Matches 73; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNVEHEVNLVVEIHRIGSKNADGKLSVKFGVLFQDDRCANLFEALVGTLLKAARRKIYV 60
 DB 6 MNVDHEVNLVVEIHRIGSKNADGKLSVKFGVLFQDDRCANLFEALVGTLLKAARRKIYV 65
 QY 61 YAGELLQGVHDDVDIVL 78
 DB 66 YPGELLQGVHDDVDIVL 83

RESULT 5

AAG58854

ID AAG58854 standard; Protein; 87 AA.

XX AAG58854;

XX 18-OCT-2000 (first entry)

XX Zea mays protein fragment SEQ ID NO: 76054.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence; corn.

OS Zea mays subsp. mays.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 16-APR-1999; 99US-0128714.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132407.

XX 05-MAY-1999; 99US-0132484.

XX 06-MAY-1999; 99US-0132485.

XX 07-MAY-1999; 99US-0132486.

XX 07-MAY-1999; 99US-0132487.

XX 11-MAY-1999; 99US-0132863.

XX 14-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143442.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

Query Match 70.4%; Score 286; DB 21; Length 83;
Best Local Similarity 70.5%; Pred. No. 4.6e-30;
Matches 55; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 1 MNVEHEVNLLVVEIHRIGSKNADGKLSYKFGVLFODDRCAALFEALVTGLKAARRKIYT 60
||| : | ||||||| : || ||||| |||||||:||:||||
DB 6 MNVDEIQLEEEIHRIGSRQDGSYKVTGVLFNDRRCANIFEALVGTFLRAAKRKIVA 65

QY 61 YAGELLQGVHDDVDIVL 78
: ||||||| |:|
DB 66 FECELLQGVHDKVEITL 83

RESULT 7
AAG58354 standard; Protein; 95 AA.
XX XX AAG58354;
XX XX
DT 18-OCT-2000 (first entry)
XX XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 75317.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
PN EP1033405-A2.
XX XX
PD 06-SEP-2000.
XX XX
PF 25-FEB-2000; 200EP-0301439.
XX XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137328.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0156559.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 23-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.

PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.
 Query Match 70.4%; Score 286; DB 21; Length 161;
 Best Local Similarity 70.5%; Pred. NO. 1.le-29;
 Matches 55; Conservative 8; Mismatches 15; Indels 0; Gaps 0;
 QY 1 MNVEHEVLLVEEIHRLGSKNADGKLSVKGVLFODDRCANLFEALVGLTKAARKRIYV 60
 Db 67 MNVDEIQKLEEEIHRLGSRQTDGYSKVTGVLFFNDDRCANIFEALVGLTKAARKRIYA 126
 QY 61 YAGELLQGVHDDVDIVL 78
 Db 127 FEGELLQGVHDKVETL 144
 RESULT 9
 ID AAY39325 standard; Protein; 92 AA.
 XX AC AAY39325;
 XX DT 01-DEC-1999 (first entry)
 XX DE PSGen13 protein.
 XX KW Progression suppressed gene; PSGen; progression elevated gene; PEGen;
 XX KW tumour; reciprocal subtraction differential RNA display; RSDD;
 XX KW differential expression; gene cloning; cancer.
 XX OS Rattus sp.
 XX PN WO9943844-A1.
 XX PD 02-SEP-1999.
 XX PF 26-FEB-1999; 99WO-US04323.
 XX PR 27-FEB-1998; 98US-0032684.
 XX PR 03-NOV-1998; 98US-0185115.
 XX PR 23-NOV-1998; 98US-0197889.
 XX PA (UYCO) UNIV COLUMBIA NEW YORK.
 XX PI Fisher PB;
 XX DR WPI; 1999-550872/46.
 XX DR N-PSDB; AAZ21517.
 XX PT Identifying nucleic acids differentially expressed between two samples,
 XX PS particularly sequences involved in tumour progression -
 XX PS Claim 40; Fig 35B; 110pp; English.
 This is the amino acid sequence of the PSGen13 protein (progression suppressed gene 13). PSGen13 has suppressed expression in progressed tumour cells. The PSGen13 nucleic acid sequence was identified using new methods for identifying nucleic acids differentially expressed between two samples. The method involves performing reciprocal subtraction differential RNA display (RSDD) between the two samples to generate two subtraction samples. The subtraction samples are amplified and compared to identify those nucleic acids that are differentially expressed. The method is used to identify and clone differentially expressed genes, particularly those with increased or reduced expression during tumour cell progression, e.g. progression suppressed genes (PSGen) and progression elevated genes (PEGen). The method reduces the complexity of the band pattern produced in conventional differential RNA display (where bands may be obscured, resulting in false positive signals) since most bands common to both samples are eliminated, allowing identification and cloning of genes displaying anticipated differential expression. RSDD requires only a single anchored primer for amplification and reamplified cDNA can be analysed by reverse Northern blotting.

```
SQ Sequence 92 AA;
Query Match 57.1%; Score 232; DB 20; Length 92;
Best Local Similarity 72.7%; Pred. No. 7.4e-23;
Matches 48; Conservative 4; Mismatches 8; Indels 6; Gaps 1;

QY 1 MNVEHVNLLVEIHRIGSKNADGKLSVFGVLFDQDRCANLFEALVGLKAAKRKIYV 60
      |||||
Db 1 MNVEHVNLLVEIHRIGSKNADGKLSVFGVLFDQDRCANLFEALVGLKAAKRKIYV 60
      |||||
QY 61 YAGELL 66
      |||
Db 57 --GRLL 60

RESULT 10
ABBI14740
ID ABBI14740 standard; Protein; 77 AA.
XX
AC ABBI14740;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polypeptide SEQ ID NO 3397.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01334.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225475.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 23-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
```


Query Match 16.0%; Score 65; DB 21; Length 324;
Best Local Similarity 29.2%; Pred. No. 4.7;
Matches 21; Conservative 11; Mismatches 30; Indels 10; Gaps 2;
QY 7 VNLVVEIHRIGSKNADGKLSVKFGVLFQDDRCANLF---EALVGTAKA-----KRR 56
|: : : | : | | | : : | : | | | : | :
Db 26 VDFIIQNGGHLDAKAADGNTALHYAALYNQPDCLKLLKGRALVGTVNEAGETALDIARK 85
QY 57 KIVTYAGELLQ 68
| : | : | : | :
Db 86 KHKCECELLEQ 97

Search completed: January 17, 2003, 02:06:36
Job time : 144 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 02:00:53 ; Search time 114 Seconds
(without alignments)
146.402 Million cell updates/sec

Title: US-09-648-310-2
Perfect score: 406
Sequence: 1 MNVEVNLVIEIHLRLSK.....AGELLQGVHDDVDIVLLQD 81

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	403	99.3	81	11 Q9CXU3	Q9CXU3 mus musculus
2	390	96.1	81	4 Q9PIF3	Q9PIF3 homo sapien
3	390	96.1	141	4 Q9POA1	Q9POA1 homo sapien
4	286	70.4	161	10 O81881	O81881 arabisidopsis
5	96.5	23.8	162	5 O62234	O62234 caenorhabdi
6	82.5	20.3	371	5 O46052	O46052 drosophila
7	81.5	20.1	183	5 Q9VZW6	Q9VZW6 drosophila
8	81.5	20.1	289	5 Q22133	Q22133 caenorhabdi
9	71	17.5	2919	12 Q85431	Q85431 rice stripe
10	69	17.0	318	16 Q98QX3	Q98QX3 mycoplasma
11	68.5	16.9	395	16 Q9PP14	Q9PP14 campylobact
12	67.5	16.6	294	5 Q8STP6	Q8STP6 encephalito
13	67.5	16.6	447	10 Q9SRM7	Q9SRM7 arabisidopsis
14	67.5	16.6	462	10 Q93W32	Q93W32 arabisidopsis
15	66	16.3	530	10 Q9SN36	Q9SN36 arabisidopsis
16	65	16.0	307	4 Q9NXH7	Q9NXH7 homo sapien

17	65	16.0	407	4	Q9NXX2	Q9NXX2 homo sapien
18	65	16.0	903	4	Q8FDY4	Q8FDY4 homo sapien
19	64.5	15.9	537	2	Q43991	Q43991 acinetobact
20	64	15.8	312	17	Q8ZTY1	Q8ZTY1 pyrobaculum
21	64	15.8	599	16	Q9WZM8	Q9WZM8 thermotoga
22	63.5	15.6	271	5	Q9XZ44	Q9XZ44 lutzomyia l
23	63.5	15.6	496	10	Q9ZQ96	Q9ZQ96 arabisidopsi
24	63	15.5	514	5	Q21458	Q21458 caenorhabdi
25	63	15.5	2535	10	Q9S240	Q9S240 arabisidopsi
26	62.5	15.4	491	10	Q9ZQ99	Q9ZQ99 arabisidopsi
27	62.5	15.4	496	10	Q9ZQ98	Q9ZQ98 arabisidopsi
28	62.5	15.4	510	16	Q98BA7	Q98BA7 rhizobium l
29	62.5	15.4	565	16	O24929	O24929 helicobacte
30	62.5	15.4	631	2	Q9XB14	Q9XB14 bacillus ce
31	62.5	15.4	986	16	Q92R86	Q92R86 rhizobium m
32	62	15.3	220	5	Q9NG93	Q9NG93 physarum po
33	62	15.3	268	16	Q8YKJ4	Q8YKJ4 anabaena sp
34	62	15.3	608	16	Q8R841	Q8R841 thermoanaer
35	61.5	15.1	491	5	Q9GV11	Q9GV11 ephydatia f
36	61.5	15.1	659	5	O96069	O96069 ciona intes
37	61.5	15.1	669	5	O96068	O96068 ciona intes
38	61.5	15.1	1010	16	Q8UG04	Q8UG04 agrobacteri
39	61	15.0	217	16	Q8XEW7	Q8XEW7 salmonella
40	61	15.0	556	16	Q9JVG1	Q9JVG1 neisseria m
41	61	15.0	627	16	O9CMX9	O9CMX9 pasteurella
42	61	15.0	635	5	Q19056	Q19056 caenorhabdi
43	61	15.0	948	12	Q8QZ10	Q8QZ10 rana tigrin
44	61	15.0	3354	5	Q8T101	Q8T101 bombyx mori
45	60.5	14.9	262	16	Q98LJ1	Q98LJ1 rhizobium l

ALIGNMENTS

RESULT 1

Q9CXU3	Q9CXU3	PRELIMINARY;	PRT;	81 AA.
ID	O9CXU3			
AC	O9CXU3			
DT	01-JUN-2001 (TRENBLrel. 17, Created)			
DT	01-JUN-2001 (TRENBLrel. 17, Last sequence update)			
DT	01-JUN-2001 (TRENBLrel. 17, Last annotation update)			
DE	3110003A17Rik protein.			
GN	3110003A17Rik.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=EMERYONIC HEAD;			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,			
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saio R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,			
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,			
RA	Hayashizaki Y.;			
RA	"Functional annotation of a full-length mouse cDNA collection."			
RL	Nature 409:685-690(2001).			
DR	EMBL; AK013984; BAB29100.1;			
DR	MGI; 1920362; 3110003A17Rik.			

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berland B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Fodor C., Gabriellian C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupsky M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weislock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003476; AAF47699.1;
 DR FlyBase: FBgn0035384; CG2113.
 SQ SEQUENCE 183 AA; 20974 MW; 293AB0EA9988A61 CRC64;

 Query Match 20.1%; Score 81.5; DB 5; Length 183;
 Best Local Similarity 31.6%; Pred. No. 0.13;
 Matches 24; Conservative 16; Mismatches 33; Indels 3; Gaps 2;

 Qy 6 EVNLLVEIHLGSKN-ADGKLSVKGVLFDQDRCANLFEALVGLTKAKRKIVTYAGE 64
 Db 77 EILQLCDLIQSGRDPIDGRKVLAFGLFETYN--NISDKLLATLLGARKYGFVDFSE 134
 Qy 65 LLLQGVHDDVDVLIQ 80
 Db 135 TLFGRRDTEPVRLLR 150

 Query Match 17.5%; Score 71; DB 12; Length 2919;
 Best Local Similarity 29.2%; Pred. No. 75;
 Matches 26; Conservative 15; Mismatches 32; Indels 16; Gaps 4;

 Qy 1 MNVEHEVNLLVEIHLGSKNADGKLSV-----KFGVLFDQDRCANLFEALVGLTKAA 53
 Db 2465 MNGEIEI---VEELNKLDKGFSHRLALVERIRVGKILGSLGYSYTKCOORIEIDG----- 2516
 Qy 54 KRRKIVTYAGELLQGVHDDVDI-VLLQD 81
 Db 2517 EGNKTHRYTGEIGWGSFDDSDVCIVVQD 2545

 RESULT 10
 Q98QX3
 ID Q98QX3 PRELIMINARY; PRT; 318 AA.
 AC Q98QX3;

OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Palmer S.;
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode *C. elegans*: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL: Z35663; CA884722.1;
 SQ SEQUENCE 289 AA; 32972 MW; 20DB67F731F03098 CRC64;

 Query Match 20.1%; Score 81.5; DB 5; Length 289;
 Best Local Similarity 33.7%; Pred. No. 0.23;
 Matches 29; Conservative 14; Mismatches 26; Indels 17; Gaps 4;

 Qy 1 MNVEHEVNLLVEIHLGSKNADGK---LSVKFGLVFDQDRCANLFP-----EALVGLTKAA 53
 Db 61 VHVCRILFLCETI-----DSNADGEPHYKVKFKLFP-----NIYSFYSGDKLVGLMIRA 110
 Qy 54 KRRKIVTYAGELLQGVHDDVDVILL 79
 Db 111 RKVGLVHFEGEMLYQKQDDEKIITML 136

 RESULT 9
 Q85431
 ID Q85431 PRELIMINARY; PRT; 2919 AA.
 AC Q85431;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE RNA polymerase.
 OS Rice stripe virus (isolate T) (RSV).
 OC Viruses; SSRNA negative-strand viruses; Tenuivirus.
 OX NCBI_TaxID=36394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T;
 RX MEDLINE=9598603; PubMed=7996149;
 RA Toriyama S., Takahashi M., Sano Y., Shimizu T.;
 RT "Nucleotide sequence of RNA 1, the largest genomic segment of rice
 RT stripe virus, the prototype of the tenuivirus.";
 RL J. Gen. Virol. 75:3569-3579(1994).
 DR EMBL: D31879; BAA06677.1;
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF02338; OTU; 1.
 DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
 DR PROSITE: PS0802; OTU; 1.
 SQ SEQUENCE 2919 AA; 336872 MW; 16FB7BC57FB12FC0 CRC64;

 Query Match 17.5%; Score 71; DB 12; Length 2919;
 Best Local Similarity 29.2%; Pred. No. 75;
 Matches 26; Conservative 15; Mismatches 32; Indels 16; Gaps 4;

 Qy 1 MNVEHEVNLLVEIHLGSKNADGKLSV-----KFGVLFDQDRCANLFEALVGLTKAA 53
 Db 2465 MNGEIEI---VEELNKLDKGFSHRLALVERIRVGKILGSLGYSYTKCOORIEIDG----- 2516
 Qy 54 KRRKIVTYAGELLQGVHDDVDI-VLLQD 81
 Db 2517 EGNKTHRYTGEIGWGSFDDSDVCIVVQD 2545

 RESULT 10
 Q98QX3
 ID Q98QX3 PRELIMINARY; PRT; 318 AA.
 AC Q98QX3;

DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DE Phosphate acetyltransferase (Phosphotransacetylase) (EC 2.3.1.8).
 GN MYPU_2370.
 OS Mycoplasma pulmonis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UAB CTIP;
 RX MEDLINE=212671165; PubMed=11353084;
 RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,
 RA Moszer I., Dywig K., Wroblewski H., Viari A., Rocha E.P.C.,
 RA Blanchard A.;
 RT "The complete genome sequence of the murine respiratory pathogen
 RT Mycoplasma pulmonis".
 RL Nucleic Acids Res. 29:2145-2153(2001).
 DR EMBL; AL445563; CAC13410.1; -;
 DR MypuList; MYPU_2370; -;
 DR InterPro; IPR002505; PTA.PTB.
 DR Pfam; PF01515; PTA.PTB.1.
 KW Transferase; Acyltransferase; Complete proteome.
 SQ SEQUENCE 318 AA; 35015 MW; ECE864C1AAB7E3EF CRC64;

 Query Match 17.0%; Score 69; DB 16; Length 318;
 Best Local Similarity 36.2%; Pred. No. 7.7;
 Matches 29; Conservative 6; Mismatches 25; Indels 20; Gaps 4;

 QY 6 EVNLLVEIHRGSKNADGKLSVKGVLFGQDR-----CANLFE-----ALVGTLLKAAK 54
 Db 44 EVSLVVE-----SKSDVKGDFNVLDQDKQYESFCQDLPSRKGSLSVQKALK 96

 QY 55 RRKIVTYAGELLQGVHDDV 74
 Db 97 TRPF--YAMMLLKKGFDDV 114

 RESULT 11
 Q9PP14
 ID Q9PP14 PRELIMINARY; PRT; 395 AA.
 AC Q9PP14;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE Putative integral membrane zinc-metalloprotease.
 GN C30723C.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
 OC Campylobacter.
 OX NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCTC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
 RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 RA Whitehead S., Barrett B.G.;
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT reveals hypervariable sequences.";
 RL Nature 403:665-668(2000).
 DR EMBL; AL139076; CAB72997.1; -;
 DR MEROPS; M48.008; -;
 DR InterPro; IPR001915; Peptidase_M48.
 DR InterPro; IPR000130; Zn_MTPetidse.
 DR Pfam; PF01435; Peptidase_M48; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 395 AA; 45319 MW; 47237D2859DEB89E CRC64;

Query Match 16.9%; Score 68.5; DB 16; Length 395;
 Best Local Similarity 26.2%; Pred. No. 12;
 Matches 21; Conservative 19; Mismatches 29; Indels 11; Gaps 3;

 QY 2 NVEHEVNLLVEE-----IHRGSKNADGKLSVKGVLFGQDRCANLFEALVGTLLKAAK 54
 Db 200 NLLKISSLMKQCGFSANGVYVIDASKRDLRLNAYFGGLFKSRVV-LFDTL---LKALN 255

 QY 55 RRKIVTYAGELLQGVHDDV 74
 Db 256 ERELLAVLGHGELGHFVHKDI 275

 RESULT 12
 Q8STP6
 ID Q8STP6 PRELIMINARY; PRT; 294 AA.
 AC Q8STP6;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DE Hypothetical protein ECU09_1300.
 GN ECU09_1300.
 OS Encephalitozoon cuniculi.
 OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
 OX NCBI_TaxID=6035;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1;
 RA Genoscope;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1;
 RX MEDLINE=21576510; PubMed=11719806;
 RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
 RA Prensier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
 RA Weissenbach J., Vivares C.P.;
 RT "Genome sequence and gene compaction of the eukaryote parasite
 RT Encephalitozoon cuniculi".
 RL Nature 414:450-453(2001).
 DR EMBL; AL590451; CAD27101.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 294 AA; 32949 MW; 01845EED48E5DF2 CRC64;

 Query Match 16.6%; Score 67.5; DB 5; Length 294;
 Best Local Similarity 29.9%; Pred. No. 10;
 Matches 29; Conservative 15; Mismatches 24; Indels 29; Gaps 5;

 QY 3 VEHEVNLLVEIHRG-----GSKNADGKLSVKGVLFGQDRCANLFEALVGTLLKAAKR-- 55
 Db 123 VFHELRIYSQIHLAVEVHGTRQGD-NINKRCGVLSDWEYIARRAEA-----LKAANKIV 177

 QY 56 -----RKIVTYAGELLQGVHDDVDIVLLQD 81
 Db 178 MKSPDPPEGLVRKI-----LIHQELLSNRDIALNKD 208

 RESULT 13
 Q9SRM7
 ID Q9SRM7 PRELIMINARY; PRT; 447 AA.
 AC Q9SRM7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DE T19F11.10 protein.
 GN T19F11.10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 19:06:08 ; Search time 3174 Seconds
(without alignments)
7151.918 Million cell updates/sec

Title: US-09-648-310-1
Perfect score: 780
Sequence: 1 ggcacgagctctctctgtcc.....gtaataaaaaaaatcatgt 780

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_ov.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_or.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pi.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pla.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	780	100.0	780	6	AX456990	Sequence
2	780	100.0	780	6	AX456994	Sequence
3	550.4	70.6	90548	2	AC117858	Rattus no
4	550.4	70.6	185634	2	AC129055	Rattus no
5	482.8	61.9	692	10	AF065991	AF065991 Mus muscu
6	471.6	60.5	294830	2	AC105605	Rattus no
7	384.2	49.3	176083	2	AC102256	Mus muscu
8	384.2	49.3	189170	2	AC102536	Mus muscu
9	335.2	43.0	835	6	AX456992	Sequence
10	335.2	43.0	835	6	AX456995	Sequence
11	334.2	42.9	65563	2	AC117670	Mus muscu
12	334.2	42.8	170882	2	AC107839	Mus muscu
13	320.4	41.1	876	17	AF116682	Homo sapi
14	283.4	35.3	743	9	BC014953	Homo sapi
15	226.2	28.0	100296	9	AL590308	Human DNA
16	206.6	26.5	552	6	AX396724	Sequence
17	155	19.9	185634	2	AC129055	Rattus no
18	109.6	14.1	516	8	AY087101	Arabidops
19	92.8	11.9	368	6	AX341854	Sequence
20	92.2	11.8	406	6	AX408040	Sequence
21	87.4	11.2	190	6	E25942	Blastocyst
22	87.4	11.2	190	6	E25963	Blastocyst
23	81.4	10.4	464	8	SL293059	S.latifolia
24	78.8	10.1	98124	8	AT11611	Arabidops
25	78.8	10.1	199536	8	ATCHRIV79	Arabidops
26	75	9.6	150446	2	AC097277	Oryza sat
27	53.6	6.9	114041	2	AP001868	Homo sapi
28	53.6	6.9	165913	9	AC079616	Homo sapi
29	53.6	6.9	242184	2	AC015899	Homo sapi
30	48.6	6.2	125544	2	AC095139	Rattus no
31	48.4	6.2	32179	3	AF047663	Caenorhab
32	48.4	6.2	252120	2	AC095803	Rattus no
33	48.4	6.2	299864	2	AC006702	Caenorhab
34	47	6.0	132288	9	AC099848	Homo sapi
35	46.8	6.0	98734	2	PFMAL1P2	Plasmodiu
36	46.4	5.9	15161	6	AX252124	Sequence
37	46.4	5.9	15161	6	AX348889	Sequence
38	45.6	5.8	175416	2	AC120291	Rattus no
39	45	5.8	6956	6	AX348658	Sequence
40	44.6	5.7	12062	1	AE007586	Clostridi
41	44.6	5.7	172390	2	AC096250	Rattus no
42	44.4	5.7	186973	2	AC113755	Rattus no
43	44.4	5.7	271313	2	AC098461	Rattus no
44	44.2	5.7	176671	2	AC115133	Rattus no
45	44.2	5.7	234112	3	PFMALP2	Plasmodiu

ALIGNMENTS

RESULT 1
AX456990

LOCUS AX456990

DEFINITION Sequence 1 from Patent WO0216419.

ACCESSION AX456990

VERSION AX456990.1 GI:21715783

KEYWORDS Rattus sp.

SOURCE Rattus sp.

ORGANISM Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE 1

AUTHORS Fisher, P.B., Kang, D.C. and Su, Z.Z.

TITLE Progression suppressed gene 13 (psgen 13) and uses thereof

AX456990 780 bp DNA linear PAT 06-JUL-2002

JOURNAL Patent: WO 0216419-A 1 28-FEB-2002;
The Trustees of Columbia University in the City of New York (US)
FEATURES Location/Qualifiers
source 1. .780
/organism="Rattus sp."
/db_xref="taxon:10118"
BASE COUNT 223 a 151 c 187 g 219 t
ORIGIN
Query Match 100.0%; Score 780; DB 6; Length 780;
Best Local Similarity 100.0%; Pred. No. 5.5e-196;
Matches 780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGAGCTCTCCTCGTCCCTCCCTCTCCACTCAGAGCTTTCTTAGCCCGAACC 60
Db 1 GGCACGAGCTCTCCTCGTCCCTCCCTCTCCACTCAGAGCTTTCTTAGCCCGAACC 60
QY 61 CTTCCTTCTCTGCTGCTCCTAGGCGGGAGCTGAGTGCAGGGTTCAGACCCA 120
Db 61 CTTCCTTCTCTGCTGCTCCTAGGCGGGAGCTGAGTGCAGGGTTCAGACCCA 120
QY 121 CGCGGCGAGCAGCTCTCAGTGAAGGAAGCAATCGGAGGCTCAGCAATGACCTGGA 180
Db 121 CGCGGCGAGCAGCTCTCAGTGAAGGAAGCAATCGGAGGCTCAGCAATGACCTGGA 180
QY 181 GCATGAGTTAACCTCCTGCTGAGGAAATTCATCGTCTGGGTTCCAAAATGCCGATGG 240
Db 181 GCATGAGTTAACCTCCTGCTGAGGAAATTCATCGTCTGGGTTCCAAAATGCCGATGG 240
QY 241 GAACTGAGTGAAGTTGGGGTCTCTTCCAGAGCAGAGATGCGCCATCTCTTTGA 300
Db 241 GAACTGAGTGAAGTTGGGGTCTCTTCCAGAGCAGAGATGCGCCATCTCTTTGA 300
QY 301 AGCGTTGGTGGGAACCTCTGAAAGCCCAACGAAGCAATGTTACGTACCGAGAGA 360
Db 301 AGCGTTGGTGGGAACCTCTGAAAGCCCAACGAAGCAATGTTACGTACCGAGAGA 360
QY 361 GCTGCTTTTGCAGAGTGTTCATGATGATGATGATGATGATGATGATGATGATGATG 420
Db 361 GCTGCTTTTGCAGAGTGTTCATGATGATGATGATGATGATGATGATGATGATGATG 420
QY 421 TTTCAGATCTGGGGTATCTGTAACCTGTAACCTGTAACCTGTAACCTGTAACCTGTA 480
Db 421 TTTCAGATCTGGGGTATCTGTAACCTGTAACCTGTAACCTGTAACCTGTAACCTGTA 480
QY 481 TTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 481 TTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 541 TTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 541 TTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 601 TCTTAAGAGGTAATCGGAGAGCTAGGCAATAAATGTTTTCAGAGGTCGGAAGAGCT 660
Db 601 TCTTAAGAGGTAATCGGAGAGCTAGGCAATAAATGTTTTCAGAGGTCGGAAGAGCT 660
QY 661 TTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Db 661 TTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
QY 721 CGAACTAAAGACCACTGCGGTGGAATAATATGTTATGTTATGTTATGTTATGTTATG 780
Db 721 CGAACTAAAGACCACTGCGGTGGAATAATATGTTATGTTATGTTATGTTATGTTATG 780
RESULT 2
AX456994
LOCUS
DEFINITION Sequence 5 from Patent WO0216419.
ACCESSION AX456994
VERSION AX456994.1 GI:21715785
KEYWORDS

SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1
AUTHORS Fisher, P. B., Kang, D. C. and Su, Z. Z.
TITLE Progression suppressed gene 13 (psgen 13) and uses thereof
JOURNAL Patent: WO 0216419-A 5 28-FEB-2002;
The Trustees of Columbia University in the City of New York (US)
FEATURES Location/Qualifiers
source 1. .780
/organism="Rattus sp."
/db_xref="taxon:10118"
BASE COUNT 223 a 151 c 187 g 219 t
ORIGIN
Query Match 100.0%; Score 780; DB 6; Length 780;
Best Local Similarity 100.0%; Pred. No. 5.5e-196;
Matches 780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGAGCTCTCCTCGTCCCTCCCTCTCCACTCAGAGCTTTCTTAGCCCGAACC 60
Db 1 GGCACGAGCTCTCCTCGTCCCTCCCTCTCCACTCAGAGCTTTCTTAGCCCGAACC 60
QY 61 CTTCCTTCTCTGCTGCTCCTAGGCGGGAGCTGAGTGCAGGGTTCAGACCCA 120
Db 61 CTTCCTTCTCTGCTGCTCCTAGGCGGGAGCTGAGTGCAGGGTTCAGACCCA 120
QY 121 CGCGGCGAGCAGCTCTCAGTGAAGGAAGCAATCGGAGGCTCAGCAATGACCTGGA 180
Db 121 CGCGGCGAGCAGCTCTCAGTGAAGGAAGCAATCGGAGGCTCAGCAATGACCTGGA 180
QY 181 GCATGAGTTAACCTCCTGCTGAGGAAATTCATCGTCTGGGTTCCAAAATGCCGATGG 240
Db 181 GCATGAGTTAACCTCCTGCTGAGGAAATTCATCGTCTGGGTTCCAAAATGCCGATGG 240
QY 241 GAACTGAGTGAAGTTGGGGTCTCTTCCAGAGCAGAGATGCGCCATCTCTTTGA 300
Db 241 GAACTGAGTGAAGTTGGGGTCTCTTCCAGAGCAGAGATGCGCCATCTCTTTGA 300
QY 301 AGCGTTGGTGGGAACCTCTGAAAGCCCAACGAAGCAATGTTACGTACCGAGAGA 360
Db 301 AGCGTTGGTGGGAACCTCTGAAAGCCCAACGAAGCAATGTTACGTACCGAGAGA 360
QY 361 GCTGCTTTTGCAGAGTGTTCATGATGATGATGATGATGATGATGATGATGATGATG 420
Db 361 GCTGCTTTTGCAGAGTGTTCATGATGATGATGATGATGATGATGATGATGATGATG 420
QY 421 TTTCAGATCTGGGGTATCTGTAACCTGTAACCTGTAACCTGTAACCTGTAACCTGTA 480
Db 421 TTTCAGATCTGGGGTATCTGTAACCTGTAACCTGTAACCTGTAACCTGTAACCTGTA 480
QY 481 TTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 481 TTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 541 TTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 541 TTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 601 TCTTAAGAGGTAATCGGAGAGCTAGGCAATAAATGTTTTCAGAGGTCGGAAGAGCT 660
Db 601 TCTTAAGAGGTAATCGGAGAGCTAGGCAATAAATGTTTTCAGAGGTCGGAAGAGCT 660
QY 661 TTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Db 661 TTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
QY 721 CGAACTAAAGACCACTGCGGTGGAATAATATGTTATGTTATGTTATGTTATGTTATG 780
Db 721 CGAACTAAAGACCACTGCGGTGGAATAATATGTTATGTTATGTTATGTTATGTTATG 780

RESULT 3

AC117858/c

LOCUS

DEFINITION

AC117858 90548 bp DNA linear HTG 18-JUL-2002
 Rattus norvegicus clone CH230-344121, *** SEQUENCING IN PROGRESS
 ***, 44 unordered pieces.

ACCESSION

AC117858

VERSION

AC117858.3

KEYWORDS

HTG: HTGS, PHASE1.

SOURCE

Norway rat.

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 90548)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,

Alsbrooks,S.B., Amarantunga,H.C., Are,J.R., Ayalew,M., Banks,T.,

Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,

Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,

Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,

Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,

Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,

Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,

Devila,M.L., Davis,C., Davy-Cartoll,L., Dederich,D.A.,

Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,

Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,

Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,

Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,

Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,

Hernandez,O., Hodgson,A., Hoques,M., Holloway,C., Hollins,B.,

Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,

Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

Karlsone,E., Kelly,S., Khan,G., King,L., Korvan,J., Kovar,C.,

Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L., Lewis,L.,

Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,

Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,

Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,

Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,

Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,

Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,

Nguyen,N., Nickerson,E., Nwokenwo,S., Ogih,M., Okwuonu,G.,

Ogungbe,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,

Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,

Rives,M., Rojas,A., Rojebokan,I., Rolfe,M., Ruiz,S., Savery,G.,

Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,

Sodergren,E., Sonaite,T., Sparks,A., Stanley,H., Stone,H.,

Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,

Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,

Usmani,K., Vasquez,L., Vera,V., Villalobos,D., Vinson,R., Wang,Q.,

Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,

Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,

Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,

Weinstock,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 90548)

Worley,K.C.

Direct Submission

Submitted (11-APR-2002)

Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 90548)

Worley,K.C.

Direct Submission

Submitted (18-JUL-2002)

Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Jul 14, 2002 this sequence version replaced gi:20162682.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GVS 344121

Center clone name: CH230-344121

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 51579 bases at least Q40

Consensus quality: 55476 bases at least Q30

Consensus quality: 58229 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 44 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1

* 1374: contig of 1374 bp in length

* 1375 1474: gap of unknown length

* 1475 3073: contig of 1599 bp in length

* 3074 3173: gap of unknown length

* 3174 4494: contig of 1321 bp in length

* 4495 4594: gap of unknown length

* 4595 6029: contig of 1435 bp in length

* 6030 6129: gap of unknown length

* 6130 7189: contig of 1060 bp in length

* 7190 7289: gap of unknown length

* 7290 8550: contig of 1261 bp in length

* 8551 8650: gap of unknown length

* 8651 10244: contig of 1594 bp in length

* 10245 10344: gap of unknown length

* 10345 11679: contig of 1335 bp in length

* 11680 11779: gap of unknown length

* 11780 13221: contig of 1442 bp in length

* 13222 13321: gap of unknown length

* 13322 14558: contig of 1237 bp in length

* 14559 14658: gap of unknown length

* 14659 16268: contig of 1610 bp in length

* 16269 16368: gap of unknown length

* 16369 17767: contig of 1399 bp in length

* 17768 17867: gap of unknown length

* 17868 18989: contig of 1122 bp in length

* 18990 19089: gap of unknown length

* 19090 20904: contig of 1815 bp in length

* 20905 21004: gap of unknown length

* 21005 22656: contig of 1652 bp in length

* 22657 22756: gap of unknown length

* 22757 24988: contig of 2232 bp in length

* 24989 25088: gap of unknown length

* 25089 26704: contig of 1616 bp in length

* 26705 26804: gap of unknown length

* 26805 28418: contig of 1614 bp in length

* 28419 28518: gap of unknown length

* 28519 30708: contig of 2090 bp in length

* 30709 32506: contig of 1798 bp in length

* 32507 32606: gap of unknown length

* 32607 34405: contig of 1799 bp in length

* 34406 34505: gap of unknown length

* 34506 36185: contig of 1679 bp in length

* 36186 36284: gap of unknown length

* 36285 37912: contig of 1528 bp in length

* 37913 40252: contig of 2340 bp in length

* 40253 40352: gap of unknown length

* 40353 42098: contig of 1746 bp in length

* 42099 42198: gap of unknown length

* 42199

QY	589	TTATTTTTTTTCTTAAGAAGGTAAATCGGAGACGTAGGCAATAAAATGTTTTCAGAGG	648
Db	76483	TTATTTTTTTTCTTAAGAAGGTAAATCGGAGACGTAGGCAATAAAATGTTTTCAGAGG	76424
QY	649	TGCGAAAAAGCTTTTGTCTTAAACCAATCTTAGTCTCTGCCACACTTGCACACTCGT	708
Db	76423	TGCGAAAAAGCTTTTGTCTTAAACCAATCTTAGTCTCTGCCACACTTGCACACTCGT	76364
QY	709	CAAACTGAGAGCGCAACTAAAGACCAACTCGGTGGGAAAAATATTATGTTTATGTAATAAA	768
Db	76363	CAAACTGAGAGCGCAACTAAAGACCAACTCGGTGGGAAAAATATTATGTTTATGTAATAAA	76304
QY	769	AAAAAATCATGT	780
Db	76303	AAAAAATCATGT	76292

RESULT 4	AC129055/c	LOCUS	DEFINITION	AC129055	185634 bp	DNA	linear	HTG 25-JUL-2002
VERSION	AC129055.1	GI:21954964						
KEYWORDS	HTG; HTGS_PHASE1.							
SOURCE	Rattus norvegicus							
ORGANISM	Rattus norvegicus							
REFERENCE	1	(bases 1 to 185634)						
AUTHORS	Muzny, D.M., Adams, C., Ali-Osman, F.R., Allen, C., Alsbrooks, S.L., Amarantunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Bunay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haviak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureishi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseghe, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, J., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Ogrunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wlarczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.							
TITLE	Direct Submission							
JOURNAL	Unpublished							

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 185634)

Worley, K.C.

Direct Submission

Submitted (25-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GNPS

Center clone name: CH230-143D2

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 124518 bases at least Q40

Consensus quality: 133171 bases at least Q30

Consensus quality: 139212 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 70 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 1024: contig of 1024 bp in length

* 1025 1124: gap of unknown length

* 1125 2174: contig of 1050 bp in length

* 2175 2274: gap of unknown length

* 2275 3296: contig of 1022 bp in length

* 3297 3396: gap of unknown length

* 3397 4410: contig of 1014 bp in length

* 4411 5532: contig of 1022 bp in length

* 5511 5532: contig of 1022 bp in length

* 5533 6632: contig of 1000 bp in length

* 6633 6732: gap of unknown length

* 6733 7799: contig of 1067 bp in length

* 7800 7899: gap of unknown length

* 7900 8966: contig of 1067 bp in length

* 8967 9066: gap of unknown length

* 9067 10100: contig of 1034 bp in length

* 10101 10200: gap of unknown length

* 10201 11897: contig of 1697 bp in length

* 11898 11997: gap of unknown length

* 11998 13230: contig of 1233 bp in length

* 13231 13330: gap of unknown length

* 13331 14783: contig of 1453 bp in length

* 14784 14883: gap of unknown length

* 14884 16203: contig of 1320 bp in length

* 16204 16303: gap of unknown length

* 16304 17354: contig of 1050 bp in length

* 17354 17433: gap of unknown length

* 17434 18784: contig of 1331 bp in length

* 18785 18884: gap of unknown length

* 18885 20009: contig of 1125 bp in length

* 20010 20109: gap of unknown length

* 20110 21365: contig of 1256 bp in length

* 21366 21465: gap of unknown length

* 21466 23426: contig of 1961 bp in length

* 23427 23526: gap of unknown length

* 23527 24788: contig of 1262 bp in length

* 24789 24888: gap of unknown length

* 24889 27141: contig of 2253 bp in length

* 27142 27241: gap of unknown length

* 27242 29343: contig of 2102 bp in length

29344 29444: gap of unknown length

29444 30711: contig of 1268 bp in length

30712 30811: gap of unknown length

30812 32025: contig of 1214 bp in length

32026 32125: gap of unknown length

32126 33758: contig of 1633 bp in length

33759 33858: gap of unknown length

33859 35685: contig of 1827 bp in length

35686 35785: gap of unknown length

35786 37296: contig of 1511 bp in length

37297 37396: gap of unknown length

37397 38612: contig of 1216 bp in length

38613 38712: gap of unknown length

38713 40182: contig of 1470 bp in length

40183 40282: gap of unknown length

40283 41698: contig of 1416 bp in length

41699 41798: gap of unknown length

41799 44026: contig of 2228 bp in length

44027 44126: gap of unknown length

44127 45966: contig of 1840 bp in length

45967 46066: gap of unknown length

46067 47123: contig of 1057 bp in length

47124 47223: gap of unknown length

47224 49001: contig of 1778 bp in length

49002 49101: gap of unknown length

49102 50956: contig of 1855 bp in length

50957 51056: gap of unknown length

51057 53229: contig of 2173 bp in length

53230 53329: gap of unknown length

53330 55625: contig of 2296 bp in length

55626 55725: gap of unknown length

55726 58308: contig of 2583 bp in length

58309 58408: gap of unknown length

58409 59889: contig of 1481 bp in length

59890 59989: gap of unknown length

59990 61974: contig of 1984 bp in length

61974 62073: gap of unknown length

62074 64204: contig of 2131 bp in length

64205 64304: gap of unknown length

64305 66825: contig of 2521 bp in length

66826 68925: gap of unknown length

68926 68716: contig of 1791 bp in length

68717 68816: gap of unknown length

68817 70918: contig of 2102 bp in length

70919 71019: gap of unknown length

71019 72938: contig of 1920 bp in length

72939 73038: gap of unknown length

73039 75043: contig of 2005 bp in length

75044 75143: gap of unknown length

75144 77763: contig of 2620 bp in length

77764 77863: gap of unknown length

77864 80351: contig of 2488 bp in length

80352 80451: gap of unknown length

80452 84227: contig of 3776 bp in length

84228 84327: gap of unknown length

84328 86532: contig of 2205 bp in length

86533 86633: gap of unknown length

86633 89992: contig of 3360 bp in length

89993 90092: gap of unknown length

90093 93245: contig of 3153 bp in length

93246 93345: gap of unknown length

93346 96897: contig of 3552 bp in length

96898 96997: gap of unknown length

96998 99360: contig of 2363 bp in length

99361 99460: gap of unknown length

99461 102048: contig of 2588 bp in length

102049 102148: gap of unknown length

102149 105689: contig of 3541 bp in length

105690 105789: gap of unknown length

105790 108898: contig of 3109 bp in length

108899 108998: gap of unknown length

108999 112575: contig of 3577 bp in length

112576 112675: gap of unknown length

[illegible]

REFERENCE AUTHORS

Rattus
1 (bases 1 to 294830)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Albrooks,S.L., Amaraturunge,H.C., Are,J.R., Ayelle,M., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karissone,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maneshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G., Orgunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,L., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 294830)
Worley,K.C.
Direct Submission
Submitted (09-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 294830)
Worley,K.C.
Direct Submission
Submitted (17-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jul 13, 2002 this sequence version replaced gi:18092828.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GNHM
Center clone name: CH230-107J10
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 199474 bases at least Q40
Consensus quality: 210759 bases at least Q30
Consensus quality: 218971 bases at least Q20

TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
NOTE: This is a 'working draft' sequence. It currently consists of 108 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
1 1158: contig of 1158 bp in length
1159: gap of unknown length
1259: contig of 1029 bp in length
2288: gap of unknown length
2387: gap of unknown length
3421: contig of 1034 bp in length
3521: gap of unknown length
4584: contig of 1043 bp in length
4564: gap of unknown length
5718: contig of 1054 bp in length
5818: gap of unknown length
6870: contig of 1052 bp in length
6871: gap of unknown length
8115: contig of 1145 bp in length
8215: gap of unknown length
9465: contig of 1250 bp in length
9565: gap of unknown length
10642: contig of 1077 bp in length
10743: gap of unknown length
12141: contig of 1399 bp in length
12241: gap of unknown length
13296: contig of 1055 bp in length
13396: gap of unknown length
14464: contig of 1068 bp in length
14584: gap of unknown length
15885: contig of 1321 bp in length
15885: gap of unknown length
17010: contig of 1025 bp in length
17110: gap of unknown length
18113: contig of 1003 bp in length
18213: gap of unknown length
19538: contig of 1325 bp in length
19638: gap of unknown length
21190: contig of 1552 bp in length
21290: gap of unknown length
22651: contig of 1361 bp in length
22751: gap of unknown length
24131: contig of 1380 bp in length
24231: gap of unknown length
25582: contig of 1351 bp in length
25682: gap of unknown length
26874: contig of 1192 bp in length
26974: gap of unknown length
28108: contig of 1134 bp in length
28208: gap of unknown length
29747: contig of 1539 bp in length
29847: gap of unknown length
30907: contig of 1060 bp in length
31007: gap of unknown length
32166: contig of 1159 bp in length
32167: gap of unknown length
33695: contig of 1429 bp in length
33795: gap of unknown length
35086: contig of 1291 bp in length
35186: gap of unknown length
36653: contig of 1467 bp in length
36753: gap of unknown length
37758: contig of 1005 bp in length
37858: gap of unknown length
39133: contig of 1275 bp in length
39233: gap of unknown length
40301: contig of 1068 bp in length
40401: gap of unknown length
41649: contig of 1248 bp in length

REFERENCE AUTHORS TITLE JOURNAL

COMMENT

* 41650	1749: gap of unknown length	41749: gap of unknown length	41750: contig of 1212 bp in length
* 41750	42961: contig of unknown length	42961: contig of unknown length	42962: contig of 1135 bp in length
* 43062	44196: contig of unknown length	44196: contig of unknown length	44197: contig of 1135 bp in length
* 44197	44296: contig of unknown length	44296: contig of unknown length	44297: contig of 1135 bp in length
* 44297	46375: contig of unknown length	46375: contig of unknown length	46376: contig of 1135 bp in length
* 46376	48246: contig of unknown length	48246: contig of unknown length	48247: contig of 1135 bp in length
* 48247	48346: contig of unknown length	48346: contig of unknown length	48347: contig of 1135 bp in length
* 48347	49828: contig of unknown length	49828: contig of unknown length	49829: contig of 1135 bp in length
* 49829	51934: contig of unknown length	51934: contig of unknown length	51935: contig of 1135 bp in length
* 51935	52035: contig of unknown length	52035: contig of unknown length	52036: contig of 1135 bp in length
* 52036	53214: contig of unknown length	53214: contig of unknown length	53215: contig of 1135 bp in length
* 53215	54540: contig of unknown length	54540: contig of unknown length	54541: contig of 1135 bp in length
* 54541	56192: contig of unknown length	56192: contig of unknown length	56193: contig of 1135 bp in length
* 56193	57788: contig of unknown length	57788: contig of unknown length	57789: contig of 1135 bp in length
* 57789	59130: contig of unknown length	59130: contig of unknown length	59131: contig of 1135 bp in length
* 59131	60422: contig of unknown length	60422: contig of unknown length	60423: contig of 1135 bp in length
* 60423	62231: contig of unknown length	62231: contig of unknown length	62232: contig of 1135 bp in length
* 62232	64211: contig of unknown length	64211: contig of unknown length	64212: contig of 1135 bp in length
* 64212	66324: contig of unknown length	66324: contig of unknown length	66325: contig of 1135 bp in length
* 66325	67965: contig of unknown length	67965: contig of unknown length	67966: contig of 1135 bp in length
* 67966	69787: contig of unknown length	69787: contig of unknown length	69788: contig of 1135 bp in length
* 69788	72512: contig of unknown length	72512: contig of unknown length	72513: contig of 1135 bp in length
* 72513	75209: contig of unknown length	75209: contig of unknown length	75210: contig of 1135 bp in length
* 75210	76907: contig of unknown length	76907: contig of unknown length	76908: contig of 1135 bp in length
* 76908	78731: contig of unknown length	78731: contig of unknown length	78732: contig of 1135 bp in length
* 78732	80962: contig of unknown length	80962: contig of unknown length	80963: contig of 1135 bp in length
Query Match	60.5%;	Score 471.6;	DB 2; Length 294830;
Best Local Similarity	96.3%;	Pred. No. 5.9e-114;	
Matches 525; Conservative	0;	Mismatches 15;	Gaps 4;
QY 229	AAATCGCGATGGGAACTGAGTGTGAAGTTGGGGTCTCTTCCAGACGACAGATGTGC	288	
DB 65641	AGATCGCGATGGGAACTGAGTGTGAAGTTGGGGTCTCTTCCAGACGACAGATGTGC	65700	
QY 289	CAATCTCTTTGAGCGTTGGTGGAACTCTGAAGCGCGAAACGAGGAAGATTGTTAC	348	
DB 65701	CAATCTCTTTGAGCGTTGGTGGAACTCTGAAGCGCGAAACGAGGAAGATTGTTAC	65760	
QY 349	GTAGCGAGGAGACTGCTTTTTCGAGGCTGTCATGATGATGATGATGATGATGATGATG	408	
DB 65761	GTAGCGAGGAGACTGCTTTTTCGAGGCTGTCATGATGATGATGATGATGATGATGATG	65820	
QY 409	AGATTATGTTGGTTGCGATCTGGGGTATCTGTTAACTGGAATTAATTAAGTAAAGG	468	
DB 65821	AGATTATGTTGGTTGCGATCTGGGGTATCTGTTAACTGGAATTAATTAAGTAAAGG	65880	
QY 469	ACAAACATGAAGTCTCTTATGATTTTATAGACCTTTTGAACAAAGGGGACTTGTG	528	
DB 65881	ACAAACATGAAGTCTCTTATGATTTTATAGACCTTTTGAACAAAGGGGACTTGTG	65940	
QY 529	AGAAGTCTGTTTATACCTTGAGCAACAAACATTAACATTAACATTAACATTAACATTA	588	
DB 65941	AGAAGTCTGTTTATACCTTGAGCAACAAACATTAACATTAACATTAACATTAACATTA	66000	

QY 589	-TTATTTTCTTAAAGAGGTAATCGGAGAGCTAGGCAATAAATGTTTTCAGAG	647	
DB 66001	TTTATTTTCTTAAAGAGGTAATCGGAGAGCTAGGCAATAAATGTTTTCAGAG	66060	
QY 648	GTGCGAAAAGCTTTTCTTAAACCACTTCTAGTCTCTGCCACACTTGACACTCCG	707	
DB 66061	GTGCGAAAAGCTTTTCTTAAACCACTTCTAGTCTCTGCCACACTTGACACTCCG	66118	
QY 708	TCAAAGTGAGAGCGAAGCAACCACTGCGTGGGAAAATATATATGTTATGTAATA	767	
DB 66119	GTCAAAGTGAGAGCGAAGCAACCACTGCGTGGGAAAATATATATGTTATGTAATA	66176	
QY 768	AAAAA 772		
DB 66177	AAAAA 66181		
RESULT 7			
AC102256			
LOCUS	AC102256	176083 bp	DNA linear HTG 21-AUG-2002
DEFINITION	Mus musculus clone RP24-191F12, WORKING DRAFT SEQUENCE, 17		
ACCESSION	AC102256		
VERSION	AC102256.2	GI:22381164	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	1 (bases 1 to 176083)		
TITLE	Birren, B., Nusbaum, C. and Lander, E.		
JOURNAL	Mus musculus, clone RP24-191F12		
AUTHORS	2 (bases 1 to 176083)		
	Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,		
	Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhaltier, B.,		
	Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,		
	Chepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,		
	Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,		
	Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,		
	Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,		
	Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,		
	Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,		
	Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,		
	Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,		
	McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,		
	Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,		
	Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,		
	Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,		
	Roman, J., Rosetti, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,		
	Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,		
	Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,		
	Topham, K., Travers, M., Travis, N., Wyman, D., Ye, W.J., Young, G.,		
	Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,		
	Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.		
	Direct Submission		
	Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome		
	Research, 320 Charles Street, Cambridge, MA 02141, USA		
	3 (bases 1 to 176083)		
	Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,		
	Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhaltier, B.,		
	Camarata, J., Chang, J., Chazaro, B., Chepel, Y., Collamore, A.,		
	Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,		
	Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,		
	Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,		
	Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,		
	Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,		
	Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,		
	McCarthy, M., McKernan, K., Meneus, L., Mihova, T., Mlenga, V.,		
	Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,		
	O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,		
	Phunkhang, P., Pierre, N., Raymond, C., Retta, K., Rise, C., Rogov, P.,		

Db	164644	AGAGTCCTGTTTATACCTTGNAGCAAAACATTACATGTAAATAAACAACCC---	164700
QY	589	TTATTTTTTTTTTCTTTAAGAAGGTAATCGGAGCAGCTAGGCAATAAAATGTTTTCAGAGG	648
		--TATTTATTTTCTTTAAGAAGGTAATCGGAAATGAGTAATGAAACATTTTGGAGG	164757
QY	649	TGGAAAAAGCTTTTGTTCCTTAAACCATTCTT-----AGTCTCGCCACACTTGACA	702
Db	164758	TGTGAAAAAGCTTTGTTCCTTTAAACCATCTTAAGACAAATTTCTACGGCAGCACTTGACA	164817
QY	703	CTCGGTCAAAAGTCAGGAGCGAACTTAAAGACCAACTGGGTGGAAAAATATTATGTTTATGT	762
Db	164818	TTCTGTCAAGACGAAAGACGAACTGCAGCAGCTGCCATGAAAATGT--TGTTTATGG	164875
QY	763	AATAAAAAAAA	773
Db	164876	AATAAAAAAAA	164886

RESULT	8
AC102536	
LOCUS	189170 bp DNA linear HTG 21-AUG-2002
DEFINITION	Mus musculus clone R23-69K22, WORKING DRAFT SEQUENCE, 31 unordered pieces.
ACCESSION	AC102536
VERSION	AC102536.2 GI:22380841
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE	ORGANISM
house mouse.	
Mus musculus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 189170)
	Birren, B., Nussbaum, C. and Lander, E.
	Mus musculus, clone RP23-69K22

Unpublished
2 (bases 1 to 189170)

REFERENCE
AUTHORS

Birtren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepeil,Y., Collangelo,M., Collins,S., Collimore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Freireira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Labocque,K.,
Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Liu,G.,
McLean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Olivier,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Orluby,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schurback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodores,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zemke,L., Zimmer,A. and Zody,M.

TITLE Zinnova, S., Lembeck, B., Zimmer, A. and Zooy, M.
JOURNAL Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 189170)

AUTHORS

Birren, N., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meirldrim, J., Meneus, L., Mihova, T., Mienga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,

O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 21, 2002 this sequence version replaced gi:17061622.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

Center project name: L189974

Center clone name: 69_K_22

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.950731

Assembly program: FILLAP; version 0.960731
Consensus quality: 181583 bases at least 040

Consensus quality: 184245 bases at least Q30

Consensus quality: 185286 bases at least Q20
Insert size: 194000; agarose-fp

Insert size: 186170; sum-of-contigs

Quality coverage: 6.8 in Q20 bases; agarose-fp
Quality coverage: 7.1 in Q20 bases; sum of conf-conf

[illegible]

* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

* 1 965: contlg of 965 bp in length

* * 966 1065: gap of 100 bp
* * 1066 1722: contig of 657 bp in length

* 1723 1822: gap of 100 bp

*	1823	2591: contlg of 769 bp in length
*	2592	2691: gap of 100 bp

* 2692 3852: contlg of 1161 bp in length
2692 2001: gap of 100 bp

* 3853 3952: gap of 100 bp
* 3853 3952: gap of 100 bp
* 3853 3952: gap of 100 bp

* 3933 3523: concy of 13/1 bp in length

* 5524 5623: gap of 100 bp

* 5624 6928: contig of 1305 bp in length
* 5020 7028: contig of 100 bp

★ 6523 7028: gap of 100 bp
7029 8370: contig of 1342 bp in length

* 8371 8470: gap of 100 bp

* 84/I 9644: contig of 11/4 bp in length

* 9645 9744: gap of 100 bp

* 9745 11173: contig of 1429 bp in length

* 11174 11273: gap of 100 bp
* 11274 12891: contig of 1618 bp in length

* 12892 12991: gap of 100 bp

* 12992 14773: contig of 1782 bp in length
* 14774 14873: gap of 100 bp

* 14774 143073: gap of 100 bp
14874 16304: contig of 1431 bp in length

* 16305 16404: gap of 100 bp
* 16405 19574: center of 3170 bp is 1000 bp

* 18403 18374: contig of 2170 bp in length
* 18575 18674: gap of 100 bp

* 18675 20391: contig of 1717 bp in length

* 20392 20491: gap of 100 bp
* 20492 23104: contig of 2613 bp in length

* 23105 23204: gap of 100 bp

* 23205 24923: contig of 1719 bp in length

*	24924	25023:	gap of	100 bp
*	25024	26589:	contig of	1566 bp in length
*	26590	26689:	gap of	100 bp
*	26690	29318:	contig of	2629 bp in length
*	29319	29418:	gap of	100 bp
*	29419	32049:	contig of	2631 bp in length
*	32050	32149:	gap of	100 bp
*	32150	34373:	contig of	2224 bp in length
*	34374	34473:	gap of	100 bp
*	34474	37088:	contig of	2615 bp in length
*	37089	37188:	gap of	100 bp
*	37189	40668:	contig of	3480 bp in length
*	40669	40768:	gap of	100 bp
*	40769	84330:	contig of	43562 bp in length
*	84331	84430:	gap of	100 bp
*	84431	88902:	contig of	4472 bp in length
*	88903	89002:	gap of	100 bp
*	89003	97203:	contig of	8201 bp in length
*	97204	97303:	gap of	100 bp
*	97304	105142:	contig of	7839 bp in length
*	105143	105242:	gap of	100 bp
*	105243	120007:	contig of	14765 bp in length
*	120008	120107:	gap of	100 bp
*	120108	132165:	contig of	12058 bp in length
*	132166	132265:	gap of	100 bp
*	132266	149030:	contig of	16765 bp in length
*	149031	149130:	gap of	100 bp
*	149131	182946:	contig of	33816 bp in length
*	182947	183046:	gap of	100 bp
*	183047	189170:	contig of	6124 bp in length.

FEATURES

```
location/Qualifiers
1. .189170
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-69K22"
/clone_lib="RPCI-23 Female Mouse BAC"
1. .965
```

misc feature

```

/clone_id=AFCI-23 female
1. .965
/note="assembly_fragment
clone_end:SP6
vector side:left"

```

misc_feature	vector_0	1066.	.1722	/note="assembly_fragment"
misc_feature	vector_1	1823.	.2591	/note="assembly_fragment"
misc_feature	vector_2	2692.	.3952	/note="assembly_fragment"
misc_feature	vector_3	3953.	.5523	/note="assembly_fragment"
misc_feature	vector_4	5624.	.6928	/note="assembly_fragment"
misc_feature	vector_5	7029.	.8370	/note="assembly_fragment"
misc_feature	vector_6	8471.	.9644	/note="assembly_fragment"
misc_feature	vector_7	9745.	.11173	/note="assembly_fragment"
misc_feature	vector_8	11274.	.12891	/note="assembly_fragment"
misc_feature	vector_9	12992.	.14773	/note="assembly_fragment"
misc_feature	vector_10	14874.	.16304	/note="assembly_fragment"
misc_feature	vector_11	16405.	.18574	/note="assembly_fragment"
misc_feature	vector_12	18675.	.20391	/note="assembly_fragment"
misc_feature	vector_13	20492.	.23104	/note="assembly_fragment"
misc_feature	vector_14	23205.	.24923	/note="assembly_fragment"
misc_feature	vector_15	25024.	.26589	/note="assembly_fragment"
misc_feature	vector_16	26690.	.29318	/note="assembly_fragment"

```
misc_feature
/note="assembly_fragment"
29419. 32049
```

Query Match	49.3%	Score 384.2;	DB 2;	Length 189170;
Best Local Similarity	85.1%	Pred. No. 8.7e-91;		
Matches 469;	Conservative	0;	Mismatches 68;	Indels 14; Gaps 3;

Qy 229 AAATGCCGATGGAAACCTGAGTGTGAAGTTTGGGGTCCTCTTCCAAGCAGCAGATGTGC 288
 + | | | | | | | | | | | | | | |
Db 142161 AGATGCTGATGGAAAAATTAACTGTGAAGTTTGGGGTCCTCTTCCAAGTGCAGATGTGC 142220

QY 289 CAATCTCTTTGAACCGTTGTTGGGAACTCTGAAGAGCCGCAAAACGAAGGAAGATTGTTAC 348
 |||||
 Db 142221 CAATCTCTTTGAACCGTTGTTAGGAACCTCTGAAGAGCTGCAAAACGAAGGAAGATTGTTAC 142280

QY 349 GTACGCAGGAGAGTCTTTTGCAAGGTTCATGATGATGATTGACATTGTATTGCTGCA 408
||||||| || |||||
Db 142281 ATACGCCGGGAACACTCTTTTGCAGGCGTTTCATGATGATGATTGACATTGTATTGCTGCA 142340

QY 409 AGATTAAATCTGGTTTGCAGATCTGGGGGTATCTGTTAAACCTGGAATTAATTAGTTAAAGG 468
|||||
D6 142341 AGATTAAATGTGTTTGCATGGCTTGGTGTATCTGATAAACCTGGAATTAAGTAAAG 142400
|||||

[illegible]

Qy 529 AGAAGTCCTGTTTTTATACCTTGGAGCAAAACATTACAATCTAAAAATAAACAAAAACCTG 588
|||||
Db 142461 AGAAGTCCTGTTTTTATACCTTGGAGCAAAACATTACAATCTAAAAATAAACAAAAAC --- 142517
|||||

Oy 589 TTTATTTTCTTCCTTAAGAAGGTAATCGGGAGACGTAGGCAATAAAATGTTTTCAGAGG 648
| | | | | | | | | | | | | | | | | |
Db 142518 ---TATTATTTTCTTCCTTAAGAAGGTAATCGGGAAAATCTAGGTAATGAACAATTTTTGGGCG 142574

Qy	649	TGCGAAAAGCTTTTGT	TTTCTTTAAACCAATCTT	-----AGTCTCGCCACACTTGACA	702
Dh	14575	TCCTGAAAAGCTTTTGT	TTTCTTTAAACCAATCTT	-----AGTCTCGCCACACTTGACA	14584

Qy 703 CTCCGTCAAAGTGAGAGCGCAACTAAAGACCAACTCGGGTGGAAAATATTATGTTTATGT 762

QY 763 AATAAAAAAA 773
|||||
142603 142703

RESULT 9
2045000

LOCUS AX456992 835 bp DNA linear PAT 06-JUL-2001
DEFINITION Sequence 3 from Patent WO0216419.
ACCESSION AX456992

PERSON 01-21/13704
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
1 Fisher, P. B., Kang, D. C. and Su, Z. Z.
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
Lakshyoda, Metazoa; Chordata; Cnidaria; Vertebrata; Euteleostomi.

FEATURES

The Trustees of Columbia University in the City of New York (US)
Location/Qualifiers

Patent: WO 0216419-A 3 28-FEB-2002;
Progression suppressed gene 13 (psgen 13) and uses thereof

```

source
l. ,835
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT      246 a   160 c   176 g   253 t

```

ORIGIN

Query Match	43.0%;	Score 335.2;	DB 6;	Length 835;
Best Local Similarity	69.1%;	Pred No. 5	6e-78;	

Matches 539; Conservative 0; Mismatches 218; Indels 23; Gaps 5;

QY 16 CGTCCCTCCCTCTCCACTGAGCGCTTCTCTAGCCGCAACCACTCTCTCTCTCTGCT 75
 DB 43 CTGTCTCTCTCTCCACAGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGCT 102
 QY 76 TGTCT 135
 DB 103 TCACCT 162
 QY 136 TTCAGTGAAGAGGAAGCAATCGGAGGCTGCAATGAACGCTGGAGCATGAGTTAACT 195
 DB 163 CCGCGGGAAGGAAGCAATCGGAGGCTGCAATGAACGCTGGAGCATGAGTTAACT 222
 QY 196 CTGGTGGAGGAATTCATCGTCTGGGTTCCAAAATGCCGATGGAACTGAGTGTGA 255
 DB 223 CTTAGTGGAGGAATTCATCGTCTGGGTTCCAAAATGCCGATGGAACTGAGTGTGA 282
 QY 256 GTTGGGCT 315
 DB 283 ATTTGGGCT 342
 QY 316 TCTGAAGCCGCAAAAGGAAGATTGTTACGTTACGAGGAGAGCTGCTTTTGAAGG 375
 DB 343 TCTTAAGCTGCAAAAGGAAGATTGTTACGTTACGAGGAGAGCTGCTTTTGAAGG 402
 QY 376 TGTTCATGATGATGTTGACATGATGTTGCAAGATTAATGTTGCTGAGATCTGGG 435
 DB 403 TGTTCATGATGATGTTGACATGATGTTGCAAGATTAATGTTGCTGAGATCTGGG 462
 QY 436 GTA-----TCTGTAACCTGGAATAATGTTAAAGGACAAACAT---G 477
 DB 463 GTACTGCCATTTTCTGTAACCTGGAT-ATAAGTGAAGAGCAAACTTTGA 521
 QY 478 AGTTCCTTATGATTTTATAGACCTTTGTAACAAAGGGGACT--TGTGAGAACTC 535
 DB 522 ACATACCTTATGATTTTATAGAACTTTGTAACAAAGGGGACTTTGTTAGAACT 581
 QY 536 CTGTTTTTATACCTTGGAGCAAAACATTAATGTAATAAATAAACAACCTGTTATTT 595
 DB 582 CTGCTCTTTTATATCTTGAAGAAATCTATGATGCTATATAAATAAATCTTAT 641
 QY 596 TTTTCTTGAAGAGTAATCGGAGAGTAGGCAATAAATGTTTTCAGAGGTCGAAA 655
 DB 642 ATTTTCTCAGAACTGTTAGGAATTCGAGCAATGAGATTTTTCGCGGGCAGGAT 701
 QY 714 TGAGAAGCAACTAAAGCAACCTGCGGTGGAATAATATATGTTATGTAATAAATAA 773
 DB 762 CAACAAGCAAACTGAAGCAACCTCCTATGAGAAATATATGTTATGTAATAAAGA 821

RESULT 10
 LOCUS AX456995
 DEFINITION Sequence 6 from Patent WO216419.
 ACCESSION AX456995
 VERSION AX456995.1 GI:21715786
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1
 AUTHOR Fisher, P.B., Kang, D.C. and Su, Z. Z.
 TITLE Progression suppressed gene 13 (psgen 13) and uses thereof
 JOURNAL Patent: WO 0216419-A 6 28-FEB-2002;
 The Trustees of Columbia University in the City of New York (US)
 FEATURES
 Location/Qualifiers
 1. .835

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 246 a 160 c 176 g 253 t
 ORIGIN
 Query Match 43.0%; Score 335.2; DB 6; Length 835;
 Best Local Similarity 69.1%; Pred. No. 5,6e-78;
 Matches 539; Conservative 0; Mismatches 218; Indels 23; Gaps 5;

QY 16 CGTCCCTCCCTCTCCACTGAGCGCTTCTCTAGCCGCAACCACTCTCTCTCTGCT 75
 DB 43 CTGTCTCTCTCTCCACAGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTGCT 102
 QY 76 TGTTCCTCTCTAGGCGGGAAGCTGAGTGCAGGTTTCAGACCCAGCGGCGAGCGCTC 135
 DB 103 TCACCT 162
 QY 136 TTCAGTGAAGAGGAAGCAATCGGAGGCTGCAATGAACGCTGGAGCATGAGTTAACT 195
 DB 163 CCGCGGGAAGGAAGCAATCGGAGGCTGCAATGAACGCTGGAGCATGAGTTAACT 222
 QY 196 CTGGTGGAGGAATTCATCGTCTGGGTTCCAAAATGCCGATGGAACTGAGTGTGA 255
 DB 223 CTTAGTGGAGGAATTCATCGTCTGGGTTCCAAAATGCCGATGGAACTGAGTGTGA 282
 QY 256 GTTGGGCT 315
 DB 283 ATTTGGGCT 342
 QY 316 TCTGAAGCCGCAAAAGGAAGATTGTTACGTTACGAGGAGAGCTGCTTTTGAAGG 375
 DB 343 TCTTAAGCTGCAAAAGGAAGATTGTTACGTTACGAGGAGAGCTGCTTTTGAAGG 402
 QY 376 TGTTCATGATGATGTTGACATGATGTTGCAAGATTAATGTTGCTGAGATCTGGG 435
 DB 403 TGTTCATGATGATGTTGACATGATGTTGCAAGATTAATGTTGCTGAGATCTGGG 462
 QY 436 GTA-----TCTGTAACCTGGAATAATGTTAAAGGACAAACAT---G 477
 DB 463 GTACTGCCATTTTCTGTAACCTGGAT-ATAAGTGAAGAGCAAACTTTGA 521
 QY 478 AGTTCCTTATGATTTTATAGACCTTTGTAACAAAGGGGACT--TGTGAGAACTC 535
 DB 522 ACATACCTTATGATTTTATAGAACTTTGTAACAAAGGGGACTTTGTTAGAACT 581
 QY 536 CTGTTTTTATACCTTGGAGCAAAACATTAATGTAATAAATAAACAACCTGTTATTT 595
 DB 582 CTGCTCTTTTATATCTTGAAGAAATCTATGATGCTATATAAATAAATCTTAT 641
 QY 596 TTTTCTTGAAGAGTAATCGGAGAGTAGGCAATAAATGTTTTCAGAGGTCGAAA 655
 DB 642 ATTTTCTCAGAACTGTTAGGAATTCGAGCAATGAGATTTTTCGCGGGCAGGAT 701
 QY 656 AAGCTTTTGTCTTAAACCACTTCTAGTCT--CTGCCACACTTGACACTCCGTCGAA 713
 DB 702 GGGAATGTTTCTTCAATAAATAATGATGATGCTATATAAATAAATAAATAA 761
 QY 714 TGAGAAGCAACTAAAGCAACCTGCGGTGGAATAATATATGTTATGTAATAAATAA 773
 DB 762 CAACAAGCAAACTGAAGCAACCTCCTATGAGAAATATATGTTATGTAATAAAGA 821

RESULT 11
 AC117670
 LOCUS Mus musculus clone RP23-383P11, LOW-PASS SEQUENCE SAMPLING.
 DEFINITION
 ACCESSION AC117670
 VERSION AC117670.1 GI:20128437
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 65563)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Mus musculus, clone RP23-383P11
Unpublished
2 (bases 1 to 65563)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A.,
Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www.seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L23864
Center clone name: 383_P_11

* NOTE: This record contains 83 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1
* 693 792: contig of 692 bp in length
* 793 1483: contig of 691 bp in length
* 1484 1583: gap of 100 bp
* 1584 2281: contig of 698 bp in length
* 2282 2381: gap of 100 bp
* 2382 3087: contig of 706 bp in length
* 3088 3187: gap of 100 bp
* 3188 3888: contig of 701 bp in length
* 3889 3988: gap of 100 bp
* 3989 4684: contig of 696 bp in length
* 4685 4784: gap of 100 bp
* 4785 5485: contig of 701 bp in length
* 5486 5585: gap of 100 bp
* 5586 6273: contig of 688 bp in length
* 6274 6373: gap of 100 bp
* 6374 7062: contig of 689 bp in length
* 7063 7162: gap of 100 bp
* 7163 7854: contig of 692 bp in length
* 7855 7954: gap of 100 bp
* 7955 8657: contig of 703 bp in length
* 8658 8757: gap of 100 bp
* 8758 9458: contig of 701 bp in length
* 9459 9558: gap of 100 bp
* 9559 10261: contig of 703 bp in length
* 10262 10361: gap of 100 bp
* 10362 11066: contig of 705 bp in length
* 11067 11166: gap of 100 bp
* 11167 11861: contig of 695 bp in length
* 11862 11961: gap of 100 bp
* 11962 12646: contig of 685 bp in length
* 12647 12746: gap of 100 bp
* 12747 13419: contig of 673 bp in length
* 13420 13519: gap of 100 bp
* 13520 14202: contig of 683 bp in length
* 14203 14302: gap of 100 bp
* 14303 14993: contig of 691 bp in length
* 14994 15093: gap of 100 bp
* 15094 15784: contig of 691 bp in length
* 15785 15884: gap of 100 bp
* 15885 16545: contig of 661 bp in length
* 16546 16645: gap of 100 bp
* 16646 17334: contig of 689 bp in length
* 17335 17434: gap of 100 bp
* 17435 18131: contig of 697 bp in length
* 18132 18231: gap of 100 bp
* 18232 18930: contig of 699 bp in length
* 18931 19030: gap of 100 bp
* 19031 19720: contig of 690 bp in length
* 19721 19820: gap of 100 bp
* 19821 20493: contig of 673 bp in length
* 20494 20593: gap of 100 bp
* 20594 21282: contig of 689 bp in length
* 21283 21382: gap of 100 bp
* 21383 22071: contig of 689 bp in length
* 22072 22171: gap of 100 bp
* 22172 22850: contig of 679 bp in length
* 22851 22950: gap of 100 bp
* 22951 23643: contig of 693 bp in length
* 23644 23743: gap of 100 bp
* 23744 24436: contig of 693 bp in length
* 24437 24536: gap of 100 bp
* 24537 25239: contig of 703 bp in length
* 25240 25339: gap of 100 bp
* 25340 26037: contig of 698 bp in length
* 26038 26137: gap of 100 bp
* 26138 26841: contig of 704 bp in length
* 26842 26941: gap of 100 bp
* 26942 27642: contig of 701 bp in length
* 27643 27742: gap of 100 bp
* 27743 28445: contig of 703 bp in length
* 28446 28545: gap of 100 bp
* 28546 29239: contig of 694 bp in length
* 29240 29339: gap of 100 bp
* 29340 30019: contig of 680 bp in length
* 30020 30119: gap of 100 bp
* 30120 30801: contig of 682 bp in length
* 30802 30901: gap of 100 bp
* 30902 31589: contig of 688 bp in length
* 31590 31689: gap of 100 bp
* 31690 32390: contig of 701 bp in length
* 32391 32490: gap of 100 bp
* 32491 33163: contig of 673 bp in length
* 33164 33263: gap of 100 bp
* 33264 33957: contig of 694 bp in length
* 33958 34057: gap of 100 bp
* 34058 34749: contig of 692 bp in length
* 34750 34849: gap of 100 bp
* 34850 35523: contig of 674 bp in length
* 35524 35623: gap of 100 bp
* 35624 36312: contig of 689 bp in length
* 36313 36412: gap of 100 bp
* 36413 37103: contig of 691 bp in length

TITLE
JOURNAL
COMMENT

[illegible]

Qy 459 AAGTTAAAGGACAACATGAAGTTCCTATGTATTTTTATATAGACTTTGGTAACAAAA 516
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 27541 AAGTTAAAGGACTAATGTGAATTCCCTATATCTATCTTTTATAGAAGCTTTTTTTATANGA 27598

RESULT 12
AC107839/c
LOCUS
DEFINITION AC107839 linear HTG 24-AUG-2002
pieces.
musculus clone RP23-284K1, WORKING DRAFT SEQUENCE, 5 ordered
fragments.

ACCESSION AC107839 GI:22474886
VERSION
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 170882)
Birken, B., Nusbaum, C., and Lander, E.
TITLE Mus musculus, clone RP23-284K1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 170882)
Birken, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campiano, A., Chang, J., Chazaro, B.,
Choepeil, F., Colangelo, M., Collins, S., Collimore, A., Cook, A.,
Cooke, P., DeArillano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagob, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
Landers, T., Lehocsky, J., Levine, R., Liu, G., Maclean, C.,
Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Rieback, M., Riley, K., Rise, C., Rogov, P., Roman, J.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A., and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 3 (bases 1 to 170882)
Birken, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepeil, Y., Collimore, A.,
Cook, A., Cooke, P., DeArillano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagob, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, C., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A., and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 24, 2002 this sequence version replaced gi:20163124.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)

```
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L20600
Center clone name: 284_K_1
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 169045 bases at least Q40
Consensus quality: 169800 bases at least Q30
Consensus quality: 170139 bases at least Q20
Insert size: 170000; agarose-fp
Insert size: 170482; sum-of-contigs
Quality coverage: 13.6 in Q20 bases; agarose-fp
Quality coverage: 13.6 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* *
*   1    107625: contig of 107625 bp in length
* *   107626 107725: gap of      100 bp
* *   107726 111441: contig of 3716 bp in length
* *   111442 111541: gap of       100 bp
* *   111542 120063: contig of 8522 bp in length
* *   120064 120163: gap of       100 bp
* *   120164 161248: contig of 41085 bp in length
* *   161249 161348: gap of       100 bp
* *   161349 170882: contig of 9534 bp in length.
* *                                     Location/Qualifiers
          source                    1..170882
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /clone="RP23-284K1"
            /clone_lib="RPCI-23 Female Mouse BAC"
misc_feature              1..107625
                           /note="assembly_fragment
vector_side:left
                            clone_end:SP6
                            107726..111441
misc_feature              /note="assembly_fragment"
                            111542..120063
misc_feature              /note="assembly_fragment"
                            120164..161248
misc_feature              /note="assembly_fragment"
                            161349..170882
                            /note="assembly_fragment
                            clone_end:T7
vector_side:right
BASE COUNT    51529 a 37012 c 37190 g 44751 t     400 others
ORIGIN
Query Match           42.8%; Score 334.2; DB 2; Length 170882;
Best Local Similarity 89.4%; Pred No. 1,6e-77;
Matches 371; Conservative         0; Mismatches 43; Indels 1; Gaps
QY  99 CTGAGTGCAGGGTTCAGACCACCGCCGCCGAGCAGCTCTTTCA GTGAAGAGGAATTCATCG 158
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  56621 CAGGGTAGCAGSGCTCAGACCCCAGCACG -AGCAGCTCTTCCTCGTAAGAAGGAACAATCG 5656
QY  159 GAGGGTGAGCAATGAACGTGGAGCATGAGGTTAACTCTCTGTGGAGGAAATTCATCGTC 218
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  56562 GAGACTGAGCAATGAATGTGGAACATGAAGTTAACCTCTCTGTGGAGGAAATTCATCGCC 5650
```


|||||
237 GAAGGAGAGTTGACATATCAGAGAGCTCTTCTGCAAGGTGTTCAATGATGTTG 296
393 ACATTGTTGCTGCAAGATTAAATGTTTGCAGATCTGGGGTA----- 438
297 ACATTATATTACTGCAAGATTAAATGTTTACATATCTTTATGACTGCCATTTTGT 356
439 -TCGGTAACTGGAAATTAAGTTAAGGACAAACAT---GAAGTTCCTTATGATTT 494
357 TCTGGTAACTGGAAT-ATAAGTGAAGAACAAACATTTTGAACATACTTAATGATTT 415
495 TTATAGACCTTTGTAACAAAGGGA--CTGTTGAGAGCTCTGTTTATACCTTGG 552
416 TTATAGAACTTTGTAAAGAGAGATTCATGTTTGAAGTCTGCTCTTTTATATC 475
553 AGCAAAACATACATGTAATAAATAAACAACCTGTTATTTTCTTCTTAAAGAGCT 612
476 TTGAAGAAATATGATGATGCTATATAAATAATCCTATATTTTCTCAGGAATCT 535
613 AATCGGAGAGTAGGCAATTAATGTTTTCAGAGTGCGAAGAGCTTTTCTTCTTA 672
536 GGTAGGAATTCAGGCAATGAGATTTTTCGGGGCAGGATGGGAATGTTTGTTCATA 595
673 AACCATCTTAGTCT-CTGCCACACTTGACACTCCGTCAGAGTGAAGCAACATAAG 730
596 AATATATAGACATTTCTATAGATATTTGACATCTCGGAAGCAACCAACATGAAG 655
731 ACCAACTCGGTGGAATAATATTTATGTAATAAATAAATAA 773
656 ACCAACTCTATGAGAAATATATGATCTTATGTAATAAAGA 698
RESULT 15
AL590308
LOCUS Human DNA sequence from clone RP11-501K14 on chromosome 6, complete
sequence.
ACCESSION AL590308
VERSION AL590308.8 GI:14596398
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 100296)
Blakey, S.
Direct Submission
Submitted (03-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerquest@sanger.ac.uk
On Jul 4, 2001 this sequence version replaced gi:14252463.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30);
an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw,
SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6

RP11-501K14 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-501K14. It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP11-501K14 is at 1 in this sequence.
The true left end of clone RP1-225E12 is at 98297 in this sequence.
The true right end of clone RP11-94L3 is at 97304 in this sequence.
FEATURES
Location/Qualifiers
1..100296
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-501K14"
/clone_lib="RPCI-11.2"
2..99
/note="MIR repeat: matches 103. .194 of consensus"
118. .199
/note="MER47 repeat: matches 2242. .2323 of consensus"
190. .265
/note="MER47 repeat: matches 2242. .2317 of consensus"
268. .346
/note="MER47 repeat: matches 1. .79 of consensus"
502. .671
/note="MIR repeat: matches 7. .185 of consensus"
1420. .1737
/note="AluJb repeat: matches 2. .310 of consensus"
2010. .2033
/note="12 copies 2 mer tt 100% conserved"
2650. .2961
/note="AluX repeat: matches 3. .306 of consensus"
5767. .5895
/note="MIR repeat: matches 21. .166 of consensus"
6062. .6159
/note="L2 repeat: matches 2587. .2709 of consensus"
6165. .6268
/note="L2 repeat: matches 2607. .2709 of consensus"
6916. .7157
/note="AluJb repeat: matches 1. .295 of consensus"
7301. .7442
/note="L2 repeat: matches 1768. .1916 of consensus"
7746. .7999
/note="AluX repeat: matches 45. .298 of consensus"
8013. .8152
/note="MIR repeat: matches 14. .151 of consensus"
8899. .8959
/note="MIR repeat: matches 194. .256 of consensus"
9209. .9497
/note="L1MB8 repeat: matches 5863. .6175 of consensus"
9503. .9698
/note="AluJb repeat: matches 90. .281 of consensus"
9703. .10428
/note="L1MB8 repeat: matches 5126. .5857 of consensus"
10429. .10759
/note="AluJb repeat: matches 1. .310 of consensus"
10760. .10773
/note="L1MB8 repeat: matches 5114. .5126 of consensus"
10897. .11050
/note="L2 repeat: matches 821. .962 of consensus"
11051. .11360
/note="AluSg repeat: matches 1. .310 of consensus"
11361. .12691
/note="L2 repeat: matches 962. .2419 of consensus"
12850. .13009
/note="AluJb repeat: matches 129. .279 of consensus"
15295. .15341
/note="MIR repeat: matches 97. .142 of consensus"
15342. .15632
/note="AluSg repeat: matches 21. .308 of consensus"
15633. .15735
/note="MIR repeat: matches 142. .250 of consensus"

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
1 (Bases 1 to 551)
REFERENCE
AUTHORS
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Bliscain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
, Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other ESTs: id05d03.xl
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brownefas.harvard.edu)
MG1:1948459 This sequence now available from the IMAGE consortium,
for clone orders contact: info@image.llnl.gov
Seq primer: -40RP from Gbco
High quality sequence stop: 429.

FEATURES
source
Location/Qualifiers
1..551
/organism="Mus musculus"
/strain="ICR"
/db_xref="taxon:10090"
/clone="IMAGE:5662133"
/clone_lib="Melton Normalized Mixed Mouse Pancreas 1
N1-MMSI"
/sex="Both for embryonic & newborn, male for adult and
adult islet"
/dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn,
adult, mixed"
/lab_host="DH10B"
/note="Vector: SPORT1; Site_1: Not I; Site_2: Sal I; Five
libraries representing E10.5/12.5 pancreatic bud, E16.5
pancreas, newborn pancreas, adult pancreas, and adult
islets of Langerhans were separately constructed using
SuperScript plasmid library kit (Life Technologies). cDNA
was made by oligo-dT priming and size-selected by column
fractionation. Libraries were amplified once on solid
support and plasmid DNA from each library was prepared
and mixed in equal amounts. The mixed library DNA was
normalized by method #4 from Bonaldo, Lennon, and Soares
1996 Genome Research 6:791-806; 0.5 microgram
single-stranded mixed library plasmid DNA was mixed with
5 micrograms PCR product representing mixed library
inserts and hybridized to an EcoT of 6. Single-stranded
(unhybridized) plasmids were isolated by hydroxyapatite
chromatography and used to make this library."
BASE COUNT 163 a 100 c 133 g 155 t
ORIGIN

Query Match	58.6%;	Score 456.8;	DB 13;	Length 551;
Best Local Similarity	92.08;	Pred. No. 2e-104;		
Matches 504;	Conservative	0;	Mismatches 42;	Indels 2;
Gaps	2;			

QY 52 CCCGAACCACTTCCTTTCTGTTGTTCCTCCCTAGGGCGCGGAAGCTCAGTGCAGGTT 111
|||||
Db 1 CCGGGGCCACTTCCTTTCTGCTGTTCCTCCCGAGGGCGCGGAAGCCGAGTACAGGTT 60
|||||

QY 112 TCAGACCACCGGCGAGCAGCTCTTCAGTAAGAAGAAAGCAATCGAGGGTGAGCAAT 171
|||||

Db 61 TCAGACCCAGGCAGC-AGCAGCTCTCCCGTGAAGAAGAAAGCAATCGAGAGTCAGCAAT 119
|||||

QY 172 GAAGTGGGACATGAGGTTAACCTTCGTGGAGGAAATTCATCGTCTGGGTTCCAAAAA 231
|||||

```

QY 232 TCCGATGGAACACTGAGTGTGAAGTTGGGTCCTCTCCAAAGACACAGATGTGCCAA 291
Db 553 TGCTGATGAAAAAATAAGTGTGAAGTTGGGTCCTCTCCAAAGATGACAGATGTGCCAA 494
QY 292 TCTCTTTGAAGGTTGGTGGGAACCTGAAAGCCGCAAAAGGAAGGAGATTCTAGCTA 351
Db 493 TCTCTTTGAAGGTTGGTGGGAACCTGAAAGCTGCAAAAGGAAGGAGATTCTAGCTA 434
QY 352 CGCAGGAGAGCTGCTTTTTCGAAGGTTTCATGATGATGTTGACATTTGATTGCTGCAAGA 411
Db 433 CGCAGGAGAGCTGCTTTTTCGAAGGTTTCATGATGATGTTGACATTTGATTGCTGCAAGA 374
QY 412 TTAATGTGTTTGCAGATCTGGGGTATCTGTAAGTAACTGGAATAATAAGTTAAAGGACA 471
Db 373 TTAATGTGTTTGCAGATCTGGGGTATCTGTAAGTAACTGGAATAATAAGTTAAAGGACT 314
QY 472 ACATGAGAGCTGCTTTTTCGAAGGTTTCATGATGATGTTGACATTTGATTGCTGCAAGA 531
Db 313 AACGGAATTTCTTAAGTATTTTATAGACTTTGTAACAAAGGGGGCTTTGTTGAGA 254
QY 532 AGTCTCTGTTTATACCTTTGGAGCAAAACATTACAAATGTAATAATAAAACCACTGTTA 591
Db 253 AGTCTCTGTTTATACCTTTGAAGCAAAACATTACAAATGTAATAATAAAACCACTGTTA 200
QY 592 TTTTCTTTTCTTGAAGAGTAACTCGGAGACGTAGGCAATAATAATGTTTTCAGAGTGC 651
Db 199 TACTATTTTCTTGAAGAGTAACTCGGAGATGTTGGAAATGTAGGTAATGAACATTTTGGAGTGT 140
QY 652 GAAAGAGCTTTGTTTCTTAACCACTCTT-----AGTCTCTGCCACACTTGACACTC 705
Db 139 GAAAGAGCTTTGTTTCTTAACCACTCTTGAAGCAATTTCTACAGGCACTTTGACATTC 80
QY 706 CGTCAAAGTGAAGAGCACTAAAGCAACCACTCGGTTGGAAATATATGTTTATGTAAT 765
Db 79 TGTCAAAGCAAGAGCAAACTGCAGACCACTGCCATGAATAATGT-TCGTTATGGAAT 22
QY 766 AAAAAAA 773
Db 21 AAAAAAA 14

```

```

RESULT 5
BG804701
LOCUS
DEFINITION
0271-10 Mouse E14.5 retina lambda ZAP II Library Mus musculus cDNA,
mRNA sequence.
ACCESSION
BG804701
VERSION
BG804701.1 GI:17951624
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 600)
AUTHORS
Mu.X., Zhao,S., Pershad,R., Hsieh,T.-F., Scarpa,A., Wang,S.W.,
White,R.A., Beremand,P.D., Thomas,T.L., Gan,L. and Klein,W.H.
TITLE
Gene expression in the developing mouse retina by EST sequencing
and microarray analysis
JOURNAL
Nucleic Acids Res. 29 (24), 4983-4993 (2001)
MEDLINE
21671825
COMMENT
Contact: Klein WH
Department of Biochemistry and Molecular Biology
University of Texas M.D. Anderson Cancer Center
Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA
Tel: 713 792 3646
Fax: 713 790 0329.
Location/Qualifiers
1..600
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="Mouse E14.5 retina lambda ZAP II Library"
/tissue_type="neural retina"
FEATURES
source

```

```

/dev_stage="embryonic day 14.5 post-fertilization"
/note="vector: pAMPI0 (Gibco); Cloned unidirectionally.
Primer: Oligo dt. RNA Isolation: cytoplasmic RNA preps
(Manniat); Cloning Technique: CUA Cloning (Clontech,
Life Technologies); Average insert size: 1.8 Kb;
Insertion site: TAGCTCCACTGAATTCGTAGTG-->. Other
information regarding entire library may be found at
http://pga.swmed.edu/Data/Libraries/microarray_cdna_librar
ies.htm."
BASE COUNT 179 a 97 c 147 g 177 t
ORIGIN

```

```

Query Match 60.2%; Score 469.4; DB 12; Length 600;
Best Local Similarity 89.0%; Pred. No. 1.4e-107;
Matches 544; Conservative 0; Mismatches 56; Indels 11; Gaps 3;
QY 62 TTCCTTCTTCTGCTGTTCCTCCCTAGGCGCGGAAGCTGAGTCAGAGGTCAGACCCAC 121
Db 1 TTCCTTCTTCTGCTGTTCCTCCCTAGGCGCGGAAGCTGAGTCAGAGGTCAGACCCAC 60
QY 122 GCGCGAGAGCTCTTCAGTGAAGGAAGCAATCGGAGGTCAGCAATGAACCTGGAG 181
Db 61 GCAGC-AGCAGCTCTTCGTTGAAGGAAGCAATCGGAG----AGCAATGAATGTGAA 115
QY 182 CATGAGGTTAACCTCTGCTGGTGGAGAAATTCATCTGCTGGGTTCCAAAATCCGATGG 241
Db 116 CATGAAGTTAACCTCTGCTGGTGGAGAAATTCATCTGCTGGGTTCCAAAATCCGATGG 175
QY 242 AAACCTGAGTGAAGTTGGGGTCTCTTCCAAAGACACAGATGTCCTTCTTCAA 301
Db 176 AATTAAGTGAAGTTGGGGTCTCTTCCAGGATGACAGATGTCCTTCTTCAA 235
QY 302 CGCTTGGTGGAACTCTGAAAGCGCAAGAGAGATTGTTACGTACGCGAGAGAG 361
Db 236 CGCTTGGTGGAACTCTGAAAGCTGCAAAAGAGAGATTGTTACATACGCGAGGAA 295
QY 362 CTGCTTTGCAAGGTTTCATGATGATGTTGACATGTTGCTGCAAGATTAAATGCTGT 421
Db 296 CTACTTTTGAAGGTTTCATGATGATGTTGACATGTTGCTGCAAGATTAAATGCTGT 355
QY 422 TTGCAGATCTGGGGTATCTGTAACCTGGAATTAATTAAGTTAAAGGCAACATGAAGT 481
Db 356 TTGCAGCTGGGTGTTATCTGTAACCTGGAATTAATTAAGTTAAAGGCAACATGAAGT 415
QY 482 TCCCTATGTTATTTATAGACTTTTAAACAAAGGGGACTTGTGAGAAGTCTCTGTTT 541
Db 416 TCCCTATGTTATTTATAGACTTTTAAACAAAGGGGCTTGTGAGAAGTCTCTGTTT 475
QY 542 TTATACCTTGGAGCAAAACATTAATCAATGTAATAATAAAACCAACCTGTTATTTT 601
Db 476 TTATACCTTGAAGCAAAACATTAATCAATGTAATAATAAAACCAACCTGTTATTTT 529
QY 602 CTTAAGAGGTAATCGGAGACGTAGGCAATAATAATGTTTTCAGAGTGCGCAAAAGCTT 661
Db 530 CTTAAGAGGTAATCGGAGAAATAGGTAATGAACAATTTTGGAGGTGTGAAAAGCTT 589
QY 662 TTGTTTCTTA 672
Db 590 TTGTTTCTTA 600

```

```

RESULT 6
BI791389
LOCUS
DEFINITION
I005403.y1 Melton Normalized Mixed Mouse Pancreas 1 NI-WMS1 Mus
musculus cDNA clone IMAGE:5662133 5' similar to FR:081881 081881
HYPOTHETICAL 18.3 KD PROTEIN. ; mRNA sequence.
ACCESSION
BI791389
VERSION
BI791389.1 GI:15819114
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

[illegible]

/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia
) with a modified polylinker; Site: 1: Not I; Site 2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5',
TCTTACCAATCTGAAGTGGGCGCCGCGGATGTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo."
BASE COUNT 182 a 93 c 145 g 173 t
ORIGIN

Query Match 58.3%; Score 455; DB 9; Length 593;
Best Local Similarity 88.5%; Pred. No. 5.8e-104;
Matches 518; Conservative 0; Mismatches 60; Indels 7; Gaps 2;
QY 99 CTGAGTGCAGGTTTACAGCCCGGCGGAGCAGCTCTTTCAGTGAAGAGGAAGCAATCG 158
D 1 CCGAGTACAGGTTTACAGCCCGGCGGAGCAGCTCTTTCAGTGAAGAGGAAGCAATCG 59
QY 159 GAGGTCAGCAATGAAGTGGAGCATGAGTTAACCTCTCTGTTGGAGGAAATTCATCGTC 218
D 60 GAGAGTCAGCAATGAAGTGGAGCATGAGTTAACCTCTCTGTTGGAGGAAATTCATCGTC 119
QY 219 TGGGTTCCAAAATGCGGATGCGGAAGTGAAGTGGGTTTGGGTCCTCTTCCAGAGCG 278
D 120 TGGGTTCCAAAATGCGGATGCGGAAGTGAAGTGGGTTTGGGTCCTCTTCCAGAGCG 179
QY 279 ACAGATGTCCCAATCTCTTTTGAAGCGTTGGTGGGAATCTGAAAGCGGCAAAACGAAGA 338
D 180 ACAGATGTCCCAATCTCTTTTGAAGCGTTGGTGGGAATCTGAAAGCGGCAAAACGAAGA 239
QY 339 ACATGTTTACAGTACCGAGCAGCTCTTTCAGAGTGTTCATGATGTTGACATTC 398
D 240 AGATGTTTACAGTACCGAGGAGCACTACTTTTGAAGGTTTTCATGATGTTGACATTC 299
QY 399 TATTGTCGAAGATTAATGTGTTTTCAGATCTGGGGTATCTGTGTAAGTGAATTAAT 458
D 300 TATTGTCGAAGATTAATGTGTTTTCAGATCTGGGGTATCTGTGTAAGTGAATTAAT 359
QY 459 AAGTTAAGGACAACATGAAGTTCCTTAATGATTTTATAGACCTTTGTAACAAAGG 518
D 360 AAGTTAAGGACAACATGAAGTTCCTTAATGATTTTATAGAACTTTGTAACAAAGG 419
QY 519 GCACCTGTTGAGAGTCCTTTTATACCTTGGACCAAAACATTAACATGTAATAA 578
D 420 GGGCTGTTGAGAGTCCTTTTATACCTTGGACCAAAACATTAACATGTAATAA --- 475
QY 579 ACAAAACCTGTTATTTTCTTAAAGAGTAACTGGGAGACGTAGGCAATAAATG 638
D 476 --TAGCAAGACCTATTATTTCTTAAAGAGTAACTGGGAGTAACTAGGTAATAA 533
QY 639 TTTTCAGAGTGGCAAAAGCTTTTCTTAAACCAATCTTA 683
D 534 TTTTGGAGGTGTGAAAAGCTTTCTTCTGTAACCACTTCTTA 578

RESULT 8
BE687862/c
LOCUS
DEFINITION
IMAGE:3465616 3' similar to TR:081881 081881 HYPOTHETICAL 18.3 KD
PROTEIN. ; mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BE687862
uw52c09.xl Soares_mammary_gland_NMLG Mus musculus cDNA clone
IMAGE:3465616 3' similar to TR:081881 081881 HYPOTHETICAL 18.3 KD
PROTEIN. ; mRNA sequence.
BE687862
BE687862.1 GI:10075486
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 669)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1384976
High quality sequence stop: 389.
FEATURES
Location/Qualifiers
1..669
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3465616"
/clone_lib="Soares_mammary_gland_NMLG"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT7T3 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 200 a 145 c 95 g 229 t
ORIGIN
Query Match 57.9%; Score 452; DB 10; Length 669;
Best Local Similarity 85.4%; Pred. No. 3.2e-103;
Matches 555; Conservative 0; Mismatches 80; Indels 15; Gaps 4;
QY 130 CAGCTCTTTCAGTGAAGAGGAGCAATCGGAGGTCAGCAATGAGCGGAGCAATGAGT 189
D 669 CAACCTCTTCCGTTTAAAGAGGA-GCAATCGGAGAGTCAGCAATGATGTGAACATGAA 611
QY 190 TAACCTCTCTGTTGAGGAAATTCATGCTGGTTCCTCAAAATGCGATGGAACCTGAG 249
D 610 TACCTCTTCTGTTGAGGAAATTCATGCTGGTTCCTCAAAATGCGATGGAACATTAAG 551
QY 250 TGTGAAGTTTGGGTCCTCTTCCAGAGCAGACAGATGTGCCAATCTCTTTGAAGCGTTGG 309
D 550 TGTGAAGTTTGGGTCCTCTTCCAGGATGACAGATGTGCCAATCTCTTTGAAGCGTTGG 491
QY 310 GGAAGTCTGAAGCCGCAAAACGAAAGAGATTGTTACGACGAGAGAGCTGCTTTT 369
D 490 AGGAAGTCTGAAGCTGCAAAACGAAAGAGATTGTTACATACGAGGAGCACTACTTTT 431
QY 370 GCAAGTGTTCATGATGATGTTGACATTGATTCGCAAGATTAATGCTGTTGTCATGG 429
D 430 GCAAGTGTTCATGATGATGTTGACATTGATTCGCAAGATTAATGCTGTTGTCATGG 371
QY 430 CTGGGGTATCTGTTAACTGGAATAATTAAGTTTAAAGGACAAACATGAAGTTCCTTATG 489
D 370 CTGCTGTCTATCTGATAAATGGAATACTAAGTTTAAAGACTACGCTGAATTCCTTATG 311
QY 490 TATTTTATAGACCTTTGTAACAAAGGAGCATTGTTGAGAGTCTCTGTTTATATACCT 549
D 310 TATTTTATAGACCTTTGTAACAAAGGAGCATTGTTGAGAGTCTCTGTTTATATACCT 251
QY 550 TGGAGCAAAACATTACATGATGTAATAAACAACAACTGTTATTTTCTTTCTTAAGAA 609
D 250 TGAAGCAAAACATTACATGATGTAATAAACAACAACTGTTATTTTCTTTCTTAAGAA 197
QY 610 GGTAATCGGAGCAGTAGGCAATAAATGTTTTCAGAGGTCGCAAAAGCTTTTCTTTTTC 669
D 196 GGTAATCGGAGCAGTAGGCAATAAATGTTTTCAGAGGTCGCAAAAGCTTTTCTTTTTC 137
QY 670 TTAACCAATCTTT-----AGTCTCTGCCACACTTGACACTCCGCTCAAGTGAGAAGCGA 723
D 670 TTAACCAATCTTT-----AGTCTCTGCCACACTTGACACTCCGCTCAAGTGAGAAGCGA 723

Db	309	CAAGGTGTCATGAATGATGTTGACATGTTATCTGCAAGATTAATGTTGGTTGCATGGC	368
Qy	431	TGGGGGTATCTGGTAAACTTGGAAATAATTAAGTTAAAGGACAAACATGAAGTTCCTTTATGT	490
Db	369	TGTGTGATCTCGATAAACTTGAATAAATTAAGTTAAAGACTAACGTGAATTCCTTATGT	428
Qy	491	ATTTTATTAGACCTTTGTAAACAAAGGGGACTTGTGAGAAGTCCCTCTTTTATACCTT	550
Db	429	ATTTTATTAGAACCTTTGTAAACAAAGGGGGCTTGTGAGAAGTCCCTCTTTTATACCTT	488
Qy	551	GGAGCAAAACATTACAATG	569
Db	489	GAAGCAAAACATTACAATG	507

RESULT	12
BQ445291	
LOCUS	UT-M-ER0-bxn-b-05-0-UI.r1 NIH_BMAP_ER0 Mus musculus cDNA clone
DEFINITION	IMAGE:5710372 5', mRNA sequence.
ACCESSION	BQ445291
VERSION	BQ445291.1
KEYWORDS	GI:21248403
SOURCE	Est.
ORGANISM	house mouse.
REFERENCE	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL	1 (bases 1 to 547)
COMMENT	NIH-MGC http://mgc.nci.nih.gov/ .
	National Institutes of Health, Mammalian Gene Collection (MGC)
	Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.
	Email: cgabbs@email.nih.gov
	Tissue Procurement: Dr. James Lin, University of Iowa
	cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
	cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
	Clone Distribution: MGC clone distribution information can be
	found through the I.M.A.G.E. Consortium/LLNL at:
	http://image.llnl.gov
	This clone was contributed by the Brain Molecular Anatomy Project
	(BMAP)
	Seq primer: pyx-5.

[illegible]

RESULT 13
 BE689324
 LOCUS BE689324 504 bp mRNA linear EST 11-SEP-2000
 DEFINITION uw52c09.y1 Soares_mammary_gland_MLMLG Mus musculus cDNA clone
 IMAGE:3465616 5' similar to TR:O81881 O81881 HYPOTHETICAL 18.3 KD
 PROTEIN, ;, mRNA sequence.
 ACCESSION BE689324
 VERSION BE689324.1 GI:10076948
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 504)
 AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Other_ESTs: uw52c09.x1
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@image.nih.gov
 This clone is available royalty-free through LML; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MG1:1384976
 Seq primer: -40RP from Gibco
 High quality sequence stop: 471.
 Location/Qualifiers
 1..504
 /organism="Mus musculus"
 FEATURES
 source

```

/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/notes="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia
) with a modified polylinker; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5].
TGTTACCAATCTCAAGTGGGCGCGCGCAATGTTTTTTTTTTTTTTTTTTT
T 3'; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

```

BASE COUNT 185 a 96 c 143 g 167 t

Query Match 54.5%; Score 425; DB 9; Length 591;
 Best Local Similarity 88.7%; Pred. No. 2.le-96;
 Matches 496; Conservative 0; Mismatches 55; Indels 8; Gaps 3;

QY 103 GTGAGGTTTCAGACCCAGCGCGAGCAGCTCTTCAGTGAAGAGGAGCAATCGGAG 162
 Db 7 GTACAGGTTTCAGACCCAGCGAGC-AGCAGCTCTTCGTGAAGAGGAGCAATCGGAG 65

QY 163 GTCAGCAATGAACGTTGAGCATGAGTTAACTCTCTGTTGAGGAGAAATTCATCGTGG 222
 Db 66 GTCAGCATGAATGTTGAACATGAGTTAACTCTCTGTTGAGGAGAAATTCATCGTGG 125

QY 223 TTCCAAATGCCGATGGGAACTGAGTGTGAAGTTTGGGTCCTCTTCCAGAGCAGAC 282
 Db 126 TTCCAAATGCTGATGAAAAATTAAGTGTGAAGTTTGGGTCCTCTTCCAGATGACAG 185

QY 283 ATGTGCCAATCTTTGAAGCGTTGGTGGGACTCTGAAGCCGCAAAACGAGGAAGAT 342
 Db 186 ATGTGCCAATCTTTGAAGCGTTGGTGAAGTGTGAAGTTTGGGTCCTCTTCCAGATGACAG 185

QY 343 TGTTACGTACGAGGAGAGCTGCTTTTCAAGGTTTTCATGATGATGTTGACATTTGATT 402
 Db 246 TGTTACATACGAGGAGAACTACITTTTCAAGGTTTTCATGATGATGTTGACATTTGATT 305

QY 403 GCTCCAGATTAATGTTGTTTGCAGATCTGGGATCTGTTAACTGGAATAAATTAAGT 462
 Db 306 GCTCCAGATTAATGTTGTTTGCATGCTGTTGATGATGATGATGATGATGATGATGAT 365

QY 463 TAAAGGACAAACATGAAGTCTCTATGATTTTATAGACCTTTCTAACAAGGGGAC 522
 Db 366 TAAAGACTACCTGAATTTCTTATGATTTTATAGACCTTTCTAACAAGGGGAC 425

QY 523 TTGTTGAGAGTCTGTTTATACCTTGGAGCAAAACATTACATGTAATAAATAACAA 582
 Db 426 TTGTTGAGAGTCTGTTTATACCTTGGAGCAAAACATTACATGTAATAAATAACAA 485

QY 583 AACCTGTTATTTTTCATGAAGAGTAAATCGGGAGCTAGGCAATAAATGTTTT 642
 Db 486 ACC-----TATTATTTTCCTAAGAGGTAATTTGGAAATGTAGTAAAT-GAATTTT 538

QY 643 CAGAGGTCGAAAGGCTT 661
 Db 539 TGGAGGTGTGAAGGCTT 557

RESULT 11
 AA008262 507 bp mRNA linear EST 25-JUL-1996
 LOCUS mg/7901.r1 Soares mouse embryo NM013.5 14.5 Mus musculus cDNA
 DEFINITION clone IMAGE:439056 5', mRNA sequence.
 AA008262
 AA008262.1 GI:1464233
 VERSION EST.
 KEYWORDS house mouse.
 SOURCE Mus musculus

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 507)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaha,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennan,G., Soares,B., Wilson,R. and
 Waterston,R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:264392

Seq primer: ETPRimer
 High quality sequence stop: 359.

Location/Qualifiers
 1..507

FEATURES
 source
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone_image="IMAGE:439056"
 /clone_lib="Soares mouse embryo NM013.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5].
 TGTTACCAATCTCAAGTGGGCGCGCGCAATGTTTTTTTTTTTTTTTTTTT
 T 3'), on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
 State Univ., from 2]; double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT73 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M.Fatima Bonaldo."

BASE COUNT 148 a 87 c 133 g 138 t 1 others
 ORIGIN

Query Match 53.9%; Score 420.4; DB 9; Length 507;
 Best Local Similarity 91.4%; Pred. No. 3e-95;
 Matches 456; Conservative 0; Mismatches 42; Indels 1; Gaps 1;

QY 71 CTGCTTGTCTCCCTAGGGCGGGAAGTGTAGTCAGGGTTTCAGACCCAGCGGAGC 130
 Db 10 CNGAGGGTTCTCTCCCGAGCGCGGAACCTGTAGTACAGGGTTTCAGACCCAGCGAGC-AGC 68

QY 131 AGCTCTTCAGTGAAGAGGAACAATCGAGGGTTCAGCAATGACGTGGAGCATGAGTT 190
 Db 69 AGCTCTTCGTGAAGAGGAACAATCGAGAGTTCAGCAATGATGTGGAACATGAAGTT 128

QY 191 AACCTCTGGTGGAGGAAATTCATCTCTGGGTTTCAAAAATCCGATGGAACATGAGT 250
 Db 129 AACCTCTGGTGGAGGAAATTCATCTCTGGGTTTCAAAAATCCGATGGAACATGAGT 188

QY 251 GTGAAGTTTGGGTCTCTTCCAGAGCAGATGTGCCAATCTCTTTGAAGGTTGGTG 310
 Db 189 GTGAAGTTTGGGTCTCTTCCAGAGTACAGATGTGCCAATCTCTTTGAAGGTTGGTA 248

QY 311 GGAATCTGAAGCCCAAAACAGAGAGATTGTTTACGTACGAGAGAGCTGCTTTTG 370
 Db 249 GGAATCTGAAGCTGCAAAACAGAGAGATTGTTTACATACGAGGGAACACTACTTTTG 308

QY 371 CAAGGTGTTTCATGATGATGTTGACATTTGATCTGCTCAAGATTAAATGTTGCCAGATC 430

```

/db_xref="taxon:10090"
/clone_lib="IMAGE:3465616"
/clone_lib="Soares_mammary_gland_NMLMG"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker: 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."
156 a 80 c 123 g 144 t 1 others

```

Query Match	53.28;	Score 414..6;	DB 10;	Length 504;
Best Local Similarity	91.6%;	Pred. No. 8.6e-94;		
Matches 460;	Conservative 0;	Mismatches 40;	Indels 2;	Gaps 2;
Qy	98	GCTGAGTGCAGGGTTACAGACCACGCGCGAGCAGCTCTTCAGTGAAGACGGAAGCAATC	157	
Db				
	1	GCCGAGTACAGGGTTACAGACCAGCGAGC-AGCAGCTCTTCCCGTGAAGAGGAAGCAATC	59	
Qy	158	GGAGGGTCAGCAATGAACGTGGAGCATGAGTCTTAACTCTGCTGGAGGAAATTCATCGT	217	
Db				
	60	GGAGAGTCAGCAATGAATGGACATGAAGTTAACTCTGCTGGAGGAAATTCATCGC	119	
Qy	218	CTGGGTTCCAAAAATGCCGATGGAAACTGAGTGTGAAGTTTGGGGTCTCTTCCAAGAC	277	
Db				
	120	CTGGGTTCCAGAAATGCTGATGGAAAAATTAAGTGTGAAGTTTGGGGTCTCTTCCAGGAT	179	
Qy	278	GACAGATGCCCAATCTCTTGAACGCTTGGTGGGAACCTCTGAAGCCGCAAAAACGGAAG	337	
Db				
	180	GACAGATGTGCCAATCTCTTGAACGCTTGGTGAAGAACTCTGAAGCTGTCAANAACGGAAG	239	
Qy	338	AAGATTGTTAGTACGACGAGGAGAGCTGCTTTGCAAGGTGTTTCATGATGATGTTGACATT	397	
Db				
	240	AAGATTGTTACATACGCAGGGGAACACTACTTTTGCAAGGTGTTTCATGATGATGTTGACATT	299	
Qy	398	GTATTTGCTGCAAGATTAATGTGGTTTTCGAGATCTGCGGGGTATCTGGTAAACTGGAATAAT	457	
Db				
	300	GTATTTGCTGCAAGATTAATGTGGTTTTCGATGGCTTGGTGTATCTGATAAACTGGAATAAC	359	
Qy	458	TAAGTTAAAGGACAAACATGAAGTTCCTTTATGTTATTTTATAGACCTTTGTAAACAAAG	517	
Db				
	360	TAAGTTAAAGACATGACGTGAATTTCCCTTATGTTATTTTATAGAACTTTGTAAACAAAG	419	
Qy	518	GGGACTTGTGTGAGAAAGTCTCTTTTTTATACCTTGGAGCAAAACATTAACAATGTAAAAATA	577	
Db				
	420	GGGCTTGTGTGAGAAAGTCTCTTTTTTATACCTTGAAGCANAACATTACAATGT-AAAAATA	478	
Qy	578	AACAAAACCTGTTATTTTTTTTTT	599	
Db				
	479	AACAAAACCTATTATTTTTTCTT	500	

RESULT 14	BF468668	501 bp	linear	EST 04-DEC-2000
LOCUS	BF468668			
DEFINITION	UI-M-BH3-atu-a-11-0-UI.r1 NIH.BMAP.M.S4 Mus musculus cDNA clone			
	UI-M-BH3-atu-a-11-0-UI 5', mRNA sequence.			
ACCESSION	BF468668			
VERSION	BF468668.1	GI:11537851		
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus			
REFERENCE	1 (bases 1 to 501)			
AUTHORS	Bonaldo, M.F., Lennon, G. and Soares, M.B.			
TITLE	Normalization and subtraction: two approaches to facilitate gene			

JOURNAL
MEDLINE
COMMENT

discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mestr@mail.nih.gov
cDNA Library Preparation: M.B. Soares Lab Clone distribution:
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
should be noted that Bento Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose
availability will be considered under appropriate and limited
collaborative arrangements
Seq primer: M13 Reverse.

```

FEATURES
source
seq primer: fwd Reverse.
Location/Qualifiers
1. .501
/organism="Mus musculus"
/dbstrain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-RH3-atu-a-11-0-UI"
/clone_lib="NIH_BMAP_M_S4"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; The
NIH_BMAP_M_S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH_BMAP_M_S4
NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
(NIH_BMAP_M_S4) was constructed as follows: PCR-amplified
cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
NIH_BMAP_M_S4 library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research
6:791-806 1996)."

```

```

BASE COUNT      124 a      115 c      123 g      139 t
ORIGIN
0:791-800, 1390]
Query Match      53.0%   Score 413.6;   DB 12;   Length 501;
Best Local Similarity 91.8%;   Pred. No. 1.5e-93;
Matches 448;   Conservative 0;   Mismatches 39;   Indels 1;   Gaps 1;
Qy      9  CTCCTCTGCTCCCTCTCTCCACTGCAGCCTTCTCTTAGCCCGGAACCACTTCCTTC 68
      |||||
Db      15  CTCCTCTGTCCCCCTCTCTCCACTGCAGCCTTCTCTTAGCCCGGGCCACTTCCTTC 74
      |||||
Qy      69  TTCTGCTGTTCCTCCCTAGGCGCGGAAGCTGAGTGCAGGTTTCAGACCACCGCGGA 128
      |||||
Db      75  TTCTGCTGTTCCTCCCGAGGCGCGGAAGCCGAGTACAGGTTTCAGACCACCGCAGC-A 133
      |||||
Qy      129  GCAGCTCTTCAGTGAAGAAAGCAATCGGAGGTCAGCAATGAACCTGGAGCATGAGG 188
      |||||
Db      134  GCAGCTCTTCGTTGAAGAAGGAAGCAATCGGAGAGTTCAGCAATGAATGTGGAACATGAAG 193
      |||||
Qy      189  TTAACTCTCTGGTGGAGGAAATTCTATCTCTGCGGTTCCAAAAATGCCCAATGGGAAACTGA 248
      |||||

```

Db	194	TTAACCTCCTGGTGGAGAAATTCATCCCTGGGTGCCAAATGCTGATGGAAATTA	253
Qy	249	GTGTGAAGTTGGGGTCCCTCTCCAAAGACACAGATGTCCAATCTCTTTGAAGCGTTGG	308
Db	254	GTGTGAAGTTGGGGTCCCTCTCCAGGATGACAGATGTCCAATCTCTTTGAAGCGTTGG	313
Qy	309	TGGGAACCTCTGAAGCGCGCAAAACGAAGGAAGATTGTACGTACGCAAGAGAGCTGCTTT	368
Db	314	TAGGAACCTCTGAAGCTGCAAAACGAAGGAAGATTGTACATACGACGAGGAACACTATT	373
Qy	369	TGCAAGTGTTTCATGATGATGTTGCACATGTTATTGCTGCAAGATTAATGTGGTTTGACA	428
Db	374	TGCAAGTGTTTCATGATGATGTTGCACATGTTATTGCTGCAAGATTAATGTGGTTTGACA	433
Qy	429	TCCTGGGGTATCTGTTAACTGGAATTAATTAAGTTAAAGGACAAACATGAAGTTCCTTAT	488
Db	434	GCITGGTGATCTGATAAAGTGGATTAAGTTAAAGACATACGTGAATTCCTTAT	493
Qy	489	GTATTTTTT 496	
Db	494	GTATTTTTT 501	

RESULT 15

BM122169/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM122169

L0506D12-3

clone L0506D12 3', mRNA sequence.

BM122169

BM122169.1

GI:17105937

EST.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 580)

Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Pantano, S., Lim, M.K. and Ko, M.S.H.

Systematic Analyses of NIA Mouse Newborn Heart cDNA Library

Unpublished (2001)

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@lgsum.grc.nia.nih.gov

Plate: L0506 row: D column: 12

Seq primer: -21M13 Forward

High quality sequence stop: 580

POLYA=Yes.

BASE COUNT	177 a	127 c	82 g	194 t	
ORIGIN					
Query Match	52.1%;	Score 406.6;	DB 13;	Length 580;	
Best Local Similarity	85.6%;	Pred. No. 8.8e-92;			
Matches 492;	Conservative 0;	Mismatches 69;	Indels 14;	Gaps 3;	
Qy 205	GGAATTCATGCTGGTGTCCAAATAATGCCGATGGAAACTCAGTGTGAAGTTGGGGT	264			
Db 580	GGAATTCATCGCTGGTGTCCAGAAATGCTGATGGAATAATTAAGTGTGAAGTTGGGT	521			
Qy 265	CCCTTTCCAAGACACAGATGTGCCAATCTTTTGAAGCGTTGGTGGGAACCTCGAAGC	324			
Db 520	CCCTTCCAGATGACAGATGTGCCAATCTTTTGAAGCGTTGGTAGGAACCTCGAAGC	461			
Qy 325	CGCAAAACGAAGGAAGATTGTTACGTACGCAGGAGAGCTGCTTTTGC AAGGTGTCATGA	384			
Db 460	TGCAAAACGAAGGAAGATTGTTACATACGCAGGGAACCTACTTTTGC AAGGTGTCATGA	401			
Qy 385	TGATGTTGACATGTATGCTGCAAGATAATGTGGTTGCAGATCTGGGGTATCTGGT	444			
Db 400	TGATGTTGACATGTATGCTGCAAGATAATGTGGTTGTCATGGCTTGGTGTATCTGAT	341			
Qy 445	AACTGGAATAATTAAGTTAAAGGACAAACATGAAGTTCCTTATGATATTTTATAGACCT	504			
Db 340	AACTGGAATAACTTAAGTTAAAGACTTAAGCTGAATTCCTTATGATATTTTATAGAACT	281			
Qy 505	TTGTAACAAAGGGGACTTGTGAGAAAGTCCTGTTTTATACCTTGGAGCAAAACATTA	564			
Db 280	TTGTAACAAAGGGGCTGTTGAGAAAGTCCTGTTTTATACCTTGAACCAAAACATTA	221			
Qy 565	CAATGTAAAAATAAACAAACCTGTTATTTTTTTTCTTTAAGAAGGTAATCGGGAGACG	624			
Db 220	CAATGTAAAAATAAACAAACCTGTTATTTTTTTTCTTTAAGAAGGTAATCGGGAAATG	167			
Qy 625	TAGCAATAAAATGTTTTTCAGAGGTGCGAAAAGCTTTTGTTCCTTAACCAATCTTT --	682			
Db 166	TAGTAATGAACCAATTTTTTGAGGTGTGAAAAGCTTTTGTCTCTTAAACCAATCTTAA	107			
Qy 683	----AGTCTCTGCCACACTTGACACTCCGTCAAAGTGAGAAAGCACTAAAGACCAACTG	738			
Db 106	GACAATTTCTACAGGCACCTTGACATCTGTCAAGCAAGGAAGCAAACTGCAGACCACTG	47			
Qy 739	CGGTGGAATAATATATCTTTATGTATATAAAAAA 773				
Db 46	CCATGAAAAATGT--TGTTTTATGAAATATAAAAAA 14				

Search completed: January 17, 2003, 02:00:47
Job time : 2236 secs

```
putative"
/codon_start=1
/protein_id="BAB29100.1"
/db_xref="GI:12851587"
/db_xref="MGI:1920362"
/translation="NMVEHVLVEEIIHRLGSNRADGKLSVKGVLFODDRCANLFE
ALVGLKAAKRRIKIVYAGELLQGVHDDVILVLD"
polyA_signal
746..751
/note="putative"
polyA_site
769
/note="putative"
BASE COUNT 228 a 146 c 174 g 221 t
ORIGIN
Query Match 69.1%; Score 539.2; DB 11; Length 769;
Best Local Similarity 86.6%; Pred. No. 3.6e-125;
Matches 658; Conservative 0; Mismatches 83; Indels 19; Gaps 5;
QY 20 CCTCCCTCTCCACGCGACCTTCTCTTAGCCGCAACCACTCTCTCTCTCTGTT 79
DB 10 CCTCCCTCTCCACGCGACCTTCTCTTAGCCGCGGCACTCTCTCTCTCTGTT 69
QY 80 CCTCCCTAGGCGCGAAGCTGAGTGCAGGTTTCAGACCCAGCGGCGAGCAGCTCTCA 139
DB 70 CCTCCCGAGGCGCGAAGCGAGTACAGGTTTCAGACCCAGCGAGC-AGCAGCTCTCC 128
QY 140 GTGAAGAAGGAAGCAATCGGAGGCTGAGCAATGAAGCTGAGCATGAGTTAACTCCTG 199
DB 129 GTGAAGAAGGAAGCAATCGGAG-AGCAATGAATGTGAACATGAAGTTAACTCCTG 184
QY 200 GTGGAGGAATATCATCTGCTGGTGTTCAGAAATGCGGATGCGGAACTGAGTGAAGTTT 259
DB 185 GTGGAGGAATATCATCTGCTGGTGTTCAGAAATGCTGAGAAATTAAGTGTGAAGTTT 244
QY 260 GGGTCTCTTCCAGACGACAGATGTGCAATCTCTTTCAGGCTGTGGGGAACCTG 319
DB 245 GGGTCTCTTCCAGATGNCAGATGTGCCAATCTCTTTCAGGCTGTGGTGAAGTCTG 304
QY 320 AAAGCGCAAAAGGAAGATGTTACGTACGACGAGAGCTGCTTTTTCGCAAGTGT 379
DB 305 AAAGCTGCAAAAGGAAGATGTTACATACGCGGGAACACTACTTTTTCGCAAGTGT 364
QY 380 CATGATGATGTCACATGTTATGCTGCAAGATTAATGCTGTTGAGATCTGGGGTAT 439
DB 365 CATGATGATGTTGACATGTTATGCTGCAAGATTAATGCTGTTGAGTCTGGTGTAT 424
QY 440 CTGGTAACTGGAATAATTAAGTTAAAGGCAACATGAAGTCTCTTATCTATTTTATA 499
DB 425 CTGATAACTGGATACTAAGTTAAAGACTACGTGAATTTCTTATGATTTTATA 484
QY 500 GACCTTTGTAACAAAGGGGACTGTTGAGAGTCTGTTTATACCTTGGAGCAAAA 559
DB 485 GAACCTTTGTAACAAAGGGGCTGTTGAGAGTCTGTTTATACCTTGAAGCAAAA 544
QY 560 CATTACAAATGAATAAACAACCTGTTATTTTCTTAAGAGGTAATCGGG 619
DB 545 CATTACAAATGAATAAACAACCTGTTATTTTCTTAAGAGGTAATTTGGG 598
QY 620 AGACGTAGGCAATAAATGTTTTCAGAGGTGCGAAAAGCTTTGTTTCTTAAACCAT 679
DB 599 AATGTAGGTATGAACATTTTTCAGAGGTGCGAAAAGCTTTGTTTCTTAAACCAT 658
QY 680 CTT-----AGTCTCTGCCACACTTGACACTCGCTCAAGTGAAGCAAGCAAGAC 733
DB 659 CTTAAGCAAAATTTCTACGACCACTTGACATCTGTCGAAGCAAGCAAGCAAGCA 718
QY 734 AACTGGGTGGAATAATTAATGTTTATGTAATAAATAA 773
DB 719 AGCTGCCATGAAAATGTT--TGTTTATGGAATAAATAA 756
```

RESULT 4
AW555045/c

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AW555045
L0249C06-3 NIA Mouse Newborn Ovary cDNA Library Mus musculus cDNA
clone L0249C06 3', mRNA sequence.
AW555045
AW555045.1 GI:7200468
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 672)
Tanaka, T.S., Jaradat, S.A., Lim, M.K., Kargul, G.J., Wang, X., Grabovac
M.J., Pantano, S., Sano, Y., Piao, Y., Nagaraja, R., Doi, H., Wood, W.H.
III, Becker, K.G. and Ko, M.S.H.
Genome-wide expression profiling of mid-gestation placenta and
embryo using a 15,000 mouse developmental cDNA microarray
Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
20381348
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdnaelgusn.grc.nia.nih.gov
Plate: L0249 row: C column: 06
Seq primer: -21M13 Forward
High quality sequence stop: 672
POLYA-Yes.

FEATURES
source

Location/Qualifiers
1..672
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="niaEST:L0249C06-3"
/db_xref="taxon:10090"
/clone="L0249C06"
/clone_lib="NIA Mouse Newborn Ovary cDNA Library"
/sex="female"
/dev_stage="Newborn Ovary"
/lab_host="DH10B"
/note="vector: pSPORT1 (Gibco/BRL Life Technology);
Site 1: SalI; Site 2: NotI; Total RNAs were extracted from
7 Newborn Ovary. The double-stranded cDNA was synthesized
by Gibco's kit with an oligo(dT) primer [NotI
primer-adaptor from Gibco/BRL]
[5'-pGACTAGTTCTAGATCGGCGCGCCCTTTTCTTTTCTTTT-3'] from
2.5ug of total RNA. The double-stranded cDNAs were
treated with T4 DNA polymerase and purified by
ethanol-precipitation. The cDNAs were ligated to
Lone-linker LL-Sal3 (include SalI sequence). The cDNAs
were purified by phenol/chloroform and separated from
free linkers by Centricon 100. Then, cDNAs were amplified
by long-range high fidelity PCR using Takara's Ex Taq
polymerase. Then, the cDNAs were purified by
phenol/chloroform and by Centricon 100. The cDNAs were
digested with SalI and NotI enzymes. Then, the cDNAs were
size selected by Gibco's Size Fractionation Column. The
cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by chemical method. The library was
constructed by Xiaohong Wang and Yulan Piao."

BASE COUNT
ORIGIN

195 a 153 c 105 g 219 t
Query Match 61.1%; Score 476.4; DB 10; Length 672;
Best Local Similarity 86.4%; Pred. No. 2.3e-109;
Matches 577; Conservative 0; Mismatches 76; Indels 15; Gaps 4;
QY 112 TCAGACCCACCGCGGAGCAGCTCTTCAGTGAAGAAGCAATCGGAGGTCAAGCAAT 171
DB 672 TCAGACCCACCGCAGC-AGCAGCTCTTCGCTGAAGAAGCAATCGGAGGTCAAGCAAT 614
QY 172 GAACGTGAGCATGAGTTTAACTTCCTGGTGGAGGAAATTCATCTCTGGTTCGCAAAA 231
DB 613 GAATGTGAGCATGAGTTTAACTTCCTGGTGGAGGAAATTCATCTCTGGTTCGCAAAA 554

Db	402	TTGATGTTGACATTGTAATGCTGCAAGATTATGTTGGTTTGCAGATCTCGGGGTATCTGGT	
QY	445	AAACTGGAATTAATTAAGTTAAAGGACAAACATGAAGTTCCTTATGTAATTTTATAGACCT	504
Db	342	AAACTGGAATTAATTAAGTTAAAGGACAAACATGAAGTTCCTTATGTAATTTTATAGACCT	283
QY	505	TTGTAAACAAAGAAGGGGACTGTTGAGAGTCTCTGTTTATACCTTGGAGGACAAACATTA	564
Db	282	TTGTAAACAAAGAAGGGGACTGTTGAGAGTCTCTGTTTATACCTTGGAGGACAAACATTA	223
QY	565	CAATGTAATAAATAACAAAACCTGTTATTTTCTTAAAGAGGTAATCGGGAGACG	624
Db	222	CAATGTAATAAATAACAAAACCTGTTATTTTCTTAAAGAGGTAATCGGGAGACG	163
QY	625	TAGGCAATAAATGTTTTCAGAGGTGCGAAAGCTTTTGTTCCTTAAACCATTCCTTAG	684
Db	162	TAGGCAATAAATGTTTTCAGAGGTGCGAAAGCTTTTGTTCCTTAAACCATTCCTTAG	103
QY	685	TCCTGCCACACTTGACACCTCCGTCAGAGTGAGAGGCACTAAAGACCAACTCGGCTGG	744
Db	102	TCCTGCCACACTTGACACCTCCGTCAGAGTGAGAGGCACTAAAGACCAACTCGGCTGG	43
QY	745	AAATATATGTTTATGTAATAAAAAAATCATGT	780
Db	42	AAATATATGTTTATGTAATAAAAAAATCATGT	7
RESULT 3			
LOCUS			
AK013984			
DEFINITION			
Mus musculus 13 days embryo head cDNA, RIKEN full-length enriched library, clone:3110003A17:homolog to PRO2013, full insert sequence.			
ACCESSION			
AK013984			
VERSION			
AK013984.1 GI:12851586			
KEYWORDS			
HTC; CAP trapper.			
SOURCE			
Mus musculus (strain:c57BL/6J) 13 days embryo head cDNA to mRNA, clone:3110003A17.			
ORGANISM			
Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE			
1 Carninci, P. and Hayashizaki, Y.			
High-efficiency full-length cDNA cloning			
Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE			
99279253			
PUBMED			
10349636			
REFERENCE			
2			
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE			
20493374			
PUBMED			
11042159			
REFERENCE			
3			
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Ozawa, Y., Izawa, M., Chata, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Tanaka, T., Matsuda, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			
RIKEN integrated sequence analysis (RISA) system-format sequencing pipeline with 384 multicapillary sequencer			
Genome Res. 10 (11), 1757-1771 (2000)			
MEDLINE			
20530913			
PUBMED			
11076861			
REFERENCE			
4			
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,			

tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the normalized osteoblast library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=yes.

FEATURES

source

Location/Qualifiers

1. .690
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone_lib="UI-R-DRI-ckz-m-14-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-DRI library is a normalized Rat Osteoblast library (nREO) constructed in pT3T7 vector according to the procedure described by Bonaldo, Lennon & Soares (Normalization and Subtraction: Two Approaches to Facilitate Gene Discovery. Genome Research 6: 791-806, 1996). The oligonucleotide used to prime first strand synthesis contained the sequence tag AAGATATCAA between the Not I cloning site and dnr18 stretch. The Rat Osteoblast tissue was provided by Lian & Stein of the University of Massachusetts Medical School.

188 a 163 c 109 g 230 t
TAG_LIB=UI-R-DRI
TAG_TISSUE=osteoblast
TAG_SEQ=AAGATATCAA

BASE COUNT

188 a 163 c 109 g 230 t

ORIGIN

Query Match 82.8%; Score 646; DB 14; Length 690;

Best Local Similarity 99.0%; Pred. No. 4.7e-152;

Matches 661; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 113 CAGACCCAGCGCGGAGCAGCTCTTCAGTGAAGAAGCAATCGGAGGTCAGCAATG 172

Db 690 CAGACCCAGCGCGGAGCAGCTCTTCAGTGAAGAAGCAATCGGAGGTCAGCAATG 631

QY 173 AAGCTGGAGATGAGGTTACCTCCCTGGTGGAGGAATTCATCGTCTGGGTTCCAAAT 232

Db 630 AAGCTGGAGATGAGGTTACCTCCCTGGTGGAGGAATTCATCGTCTGGGTTCCAAAT 571

QY 233 GCGATGGGAACTGAGTGTGAAGTTGGGTCCTCTTCCAGACGACATGTGCCAAT 292

Db 570 GCGATGGGAACTGAGTGTGAAGTTGGGTCCTCTTCCAGACGACATGTGCCAAT 513

QY 293 CTCCTTTGAAGCGTTGGTGGAACTCTGAAGCGGCAAAAGCAAGAGATTGTACGTAC 352

Db 512 CTCCTTTGAAGCGTTGGTGGAACTCTGAAGCGGCAAAAGCAAGAGATTGTACGTAC 453

QY 353 GCAGGAGAGCTGCTTTTGAAGGTTCATGATGATGTTGACATGTTGATGCTGCAAGAT 412

Db 452 GCAGGAGAGCTGCTTTTGAAGGTTCATGATGATGTTGACATGTTGATGCTGCAAGAT 393

QY 413 TAATGTTGTTGAGATCTGGGGTATCGTTAACTGGGAATTAATTAAGTTAAGGACAA 472

Db 392 TAATGTTGTTGAGATCTGGGGTATCGTTAACTGGGAATTAATTAAGTTAAGGACAA 333

QY 473 ACATGAAGTTCCTTATGATTTTATAGACCTTTGTAAACAAAGGGGACTTGTGTAGAA 532

Db 332 ACATGAAGTTCCTTATGATTTTATAGACCTTTGTAAACAAAGGGGACTTGTGTAGAA 273

QY 533 GTCCCTGTTTTATACCTGGAGCAAAACATATCAATGTAAAAATAACAAAACCTGTTAT 592

Db 272 GTCCCTGTTTTATACCTGGAGCAAAACATATCAATGTAAAAATAACAAAACCTGTTAT 213

QY 593 TTTTCTTTTCTTAAGAGGTAATCGGGAGACGTAGGCAATAAATGTTTTCAGAGGTGCG 652

Db 212 TTTTCTTTTCTTAAGAGGTAATCGGGAGACGTAGGCAATAAATGTTTTCAGAGGTGCG 153

QY 653 AAAAAGCTTTTCTTTTAAACCATCTTACTCTGCGCACACTTGACACATCCGTCAAA 712
Db 152 AAAAAGCTTTTCTTTTAAACCATCTTACTCTGCGCACACTTGACACATCCGTCAAA 93
QY 713 GTGAGAAGCGAACTAAAGACCAACTGCGGTGGAAAATATTATGTTTATATAAAAAA 772
Db 92 GTGAGAAGCGAACTAAAGACCAACTGCGGTGGAAAATATTATGTTTATATAAAAAA 33

QY 773 AATCATGT 780

Db 32 AATCATGT 25

RESULT 2

AA891725/c

LOCUS

DEFINITION

AA891725

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Apr 3, 1998 this sequence version replaced gi:3018604.

The Institute for Genomic Research

9712, Medical Center Drive, Rockville, MD 20850, USA

Tel: (301)-838-3529

Fax: (301)-838-0208

Email: nhleett@igrr.org

Seq primer: M13-21.

Location/Qualifiers

1. .642

/organism="Rattus sp."

/db_xref="taxon:10116"

/clone="RKIAG02"

/clone_lib="Normalized rat kidney, Bento Soares"

/note="Organ: kidney; Vector: pT73Pac; Site_1: EcoRI;

Site_2: NotI"

BASE COUNT

184 a 155 c 99 g 204 t

ORIGIN

Query Match 81.5%; Score 636; DB 9; Length 642;

Best Local Similarity 100.0%; Pred. No. 1.6e-149;

Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 GAAGAACCAATCGAGGTCAGCATGAACTCGAGCATGAGTTAACTCTCTCTGGTGA 204

Db 642 GAAGAACCAATCGAGGTCAGCATGAACTCGAGCATGAGTTAACTCTCTCTGGTGA 583

QY 205 GGAATTCATCGTCTGGGTTCCAAATATGCCGATGGAACTGAGTGTGAAGTTTGGGT 264

Db 582 GGAATTCATCGTCTGGGTTCCAAATATGCCGATGGAACTGAGTGTGAAGTTTGGGT 523

QY 265 CCTCTTCAAGACGACAGATGTGCAATCTCTTTGAAGCGTTGGGGAACCTCTGAAAGC 324

Db 522 CCTCTTCAAGACGACAGATGTGCAATCTCTTTGAAGCGTTGGGGAACCTCTGAAAGC 463

QY 325 CGCAAAACGAGGAGATTCTTACCTAGCAGGAGAGCTGCTTTTGCAGAGTTCATGA 384

Db 462 CGCAAAACGAGGAGATTCTTACCTAGCAGGAGAGCTGCTTTTGCAGAGTTCATGA 403

QY 385 TGATGTTGACATTGTTATGCTGCAAGATTATGTTGTTGACAGATCTGGGGTATCTGTT 444

QY

shop wodon

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 21:37:23 ; Search time 2228 seconds

(without alignments)
5669.874 Million cell updates/sec

Title: US-09-648-310-1

Perfect score: 780

Sequence: 1 ggcacgagctctctctctcc.....gtaataaaaaaaatcatgt 780

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	646	82.8	690	14	BQ192542
c 2	636	81.5	642	9	AA891725
c 3	539.2	69.1	769	11	AK013984
c 4	476.4	61.1	672	10	AW555045
c 5	469.4	60.2	600	12	BG804701
c 6	456.8	58.6	551	13	BI791389

7	455	58.3	593	9	AA544950	vk38e06.r
8	452	57.9	669	10	BE687862	BE687862 uw52c09.x
9	438.8	56.3	613	9	AA607954	AA607954 vm41a03.r
10	425	54.5	591	9	AA542157	AA542157 vj80d03.r
11	420.4	53.9	507	9	AA008262	AA008262 mg77901.r
12	415.6	53.3	547	14	BQ445291	BQ445291 UI-M-ERO-
13	414.6	53.2	504	10	BE689324	BE689324 uw52c09.y
14	413.6	53.0	501	12	BF468668	BF468668 UI-M-BH3-
15	406.6	52.1	580	13	BM122169	BM122169 L0506D12-
16	402.6	51.6	481	9	AI154666	AI154666 ud37d02.r
17	400.6	51.4	511	10	AW823811	AW823811 uf59h02.y
18	392	50.3	471	9	AI786148	AI786148 uf1e09.y
19	390	50.0	465	9	AA220360	AA220360 mv70d02.r
20	387.2	49.6	564	13	BI791071	BI791071 id05d03.x
21	381.4	48.9	457	10	BB849249	BB849249 BB849249
22	373.6	47.9	430	9	AA816139	AA816139 vP4Ac07.r
23	372.4	47.7	501	9	AA986660	AA986660 uc82e12.y
24	370	47.4	490	9	AA833425	AA833425 ub58d03.r
25	368.4	47.2	437	9	AA462630	AA462630 vf90f10.r
26	357.6	45.8	564	9	AA712021	AA712021 vu60d08.r
27	347.8	44.6	521	9	AA986390	AA986390 uc82e12.x
28	337.8	43.3	347	9	AI234233	AI234233 EST230921
29	335.2	43.0	845	11	AF161398	AF161398 Homo sapi
30	331	42.4	384	13	BM390524	BM390524 UI-R-CN1-
31	320.4	41.1	876	11	AF116682	AF116682 Homo sapi
32	314.8	40.4	797	13	BI859694	BI859694 603386283
33	313.2	40.2	425	9	AA986641	AA986641 uc82d03.y
34	308.8	39.6	365	10	BB844309	BB844309 BB844309
35	301	38.6	749	14	BQ015110	BQ015110 UI-H-ED1-
36	299.4	38.4	755	12	BG330613	BG330613 602560008
37	294.8	37.8	573	13	BI110495	BI110495 602896473
38	294.8	37.8	791	12	BG528919	BG528919 602579350
39	293.8	37.7	752	10	AV716363	AV716363 AV716363
40	293.4	37.6	747	12	BG502291	BG502291 602550569
41	291	37.3	304	14	H34607	H34607 EST111677 R
42	288	36.9	634	9	AI133406	AI133406 HA2013 Hu
43	287.6	36.9	462	10	AW492391	AW492391 UI-M-BH3-
44	287.6	36.9	884	12	BG164526	BG164526 602342067
45	286.8	36.8	678	10	AW027019	AW027019 wt71b02.x

ALIGNMENTS

RESULT 1	BQ192542/c	690 bp	linear	EST 30-APR-2002
LOCUS	UI-R-DRI-ckz-m-14-0-UI.sl	UI-R-DRI	Rattus norvegicus	CDNA clone
DEFINITION	UI-R-DRI-ckz-m-14-0-UI 3', mRNA sequence.			
ACCESSION	BQ192542			
VERSION	BQ192542.1	GI:20368093		
KEYWORDS	EST.			
SOURCE	Norway rat.			
ORGANISM	Rattus norvegicus			
REFERENCE	1 (bases 1 to 690)			
AUTHORS	Bonaldi,M.F., Lennon,G. and Soares,M.B.			
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery			
JOURNAL	Genome Res. 6 (9), 791-806 (1996)			
MEDLINE	97044477			
COMMENT	Contact: Soares, MB Program for Rat Gene Discovery and Mapping University of Iowa 451 Eckstein Medical Research Building Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: msoaresblue.weeg.uiowa.edu The sequence contained an oligo-dr track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A			

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 14.1%; Score 109.6; DB 21; Length 657;
Best Local Similarity 67.5%; Pred. No. 1.1e-20;
Matches 154; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY	169	AATGAACGTGGAGCATGAGGTTAACCTCCTGGTGGAGGAAATTCATCGTCTGGGTTCCAA	228
DB	198	AATCAACGTAGACGACGAGAGATTACAAACTGGAAGAAGAGATCCATCGTCTGGTCTCG	257
QY	229	AAATGCCGATGGGAACTGAGTGTGAAGTTGGGGTCTCTTCCAGACGACAGATGTGC	288
DB	258	TCAGACCGATGGGCTCTTACAAAGGTGACGTTGGAGTGTGTTCATGACGATCGATGTGC	317
QY	289	CAATCTCTTTGAAGCGTTGGTGGGAACTCTGAAAGCCGCAACGAAAGGAGATTTGTAC	348
DB	318	CAACATCTTTGAAGCATTTGGTGGGACTCTGAGAGCTGCCAAGAAACGCAAAATAGTCGC	377
QY	349	GTACGACGAGAGCTGCTTTTGAAGGTGTTCATGATGATGTTGACAT	396
DB	378	ATTTGAAGGTGAACCTTTTGCTTCAAGGCGTTCAGGATAAGGTTGAGAT	425

Search completed: January 16, 2003, 21:51:13
Job time : 307 secs

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141642.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0143390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158569.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.

PR	10-AUG-1999;	99US-01481171.
PR	11-AUG-1999;	99US-01483119.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150366.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-01553486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

```

-Query Match      14.7%   Score 114.6;   DB 21;   Length 551;
Best Local Similarity 67.2;   Pred. No. 3.8e-22;
Matches 162;   Conservative 0;   Mismatches 79;   Indels 0;   Gaps 0;

QY 167 GCATGACGTGGAGCATGAGGTAACTCTCGTGGAGGAATAATCAATCGTCTGGGTTC 226
|| ||||| ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db	146	GCATGACGTGGAGGAGGAGGTCTGGGAGGCTCAAGGAGGAGATCCATAGGCTCGGCCAG	201
QY	227	AAAAATGCCGATGGGAAACTCAGTGTGAAGTTTGGGTCTCTTCCAAAGACGACACATGT	286
Db	206	CAGCAGCCGATGGCTCTTACAAGGTCAAGTTTGGCGTTCTCTTCAACGACGATCGGTGT	265
QY	287	GCCATCTCTTTTGAAGCCGTTGGTGGGAACCTCTGAAGCCGCAAAACGAAGGAAGATTGTT	346
Db	266	GCAAAATATCTTTGAAGCACTAGTTGGCACCTTGAGGCCGCCCAAGAGAGAGATCTTG	325
QY	347	AGCTAGCCGAGCAGAGCTCTTTTTCGAAGGTCTTCATCATGATGTTGACATTGTATTCGTC	406
Db	326	ACCTACGAAGCGAGCTCTTCTGCAAGGTGTCCATGACAACGTGGAGATAACCCCTGTTG	385
QY	407	C 407	
Db	386	C 386	
RESULT	15		
AAC	53673		
ID	AAC53673	standard; DNA; 657 bp.	
XX	XX		
AC	AAC53673;		
XX	XX		
DT	18-OCT-2000	(first entry)	
XX	XX		
DE	XX		
XX	XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 75315.	
KW	XX	Hybridisation assay; genetic mapping; gene expression control;	
KW	XX	protein identification; signal transduction pathway;	
KW	XX	metabolic pathway; promoter; termination sequence; ss.	
OS	XX	Arabidopsis thaliana.	
XX	XX		
PN	EP1033405-A2.		
XX	XX		
PD	06-SEP-2000.		
XX	XX		
PF	25-FEB-2000; 2000EP-0301439.		
XX	XX		
PR	25-FEB-1999;	99US-0121825.	
PR	05-MAR-1999;	99US-0123180.	
PR	09-MAR-1999;	99US-0123348.	
PR	23-MAR-1999;	99US-0125788.	
PR	25-MAR-1999;	99US-0126264.	
PR	29-MAR-1999;	99US-0126785.	
PR	01-APR-1999;	99US-0127462.	
PR	06-APR-1999;	99US-0128234.	
PR	08-APR-1999;	99US-0128714.	
PR	16-APR-1999;	99US-0129845.	
PR	19-APR-1999;	99US-0130077.	
PR	21-APR-1999;	99US-0130449.	
PR	23-APR-1999;	99US-0130510.	
PR	23-APR-1999;	99US-0130891.	
PR	28-APR-1999;	99US-0131449.	
PR	30-APR-1999;	99US-0132048.	
PR	30-APR-1999;	99US-0132407.	
PR	04-MAY-1999;	99US-0132484.	
PR	05-MAY-1999;	99US-0132485.	
PR	06-MAY-1999;	99US-0132486.	
PR	06-MAY-1999;	99US-0132487.	
PR	07-MAY-1999;	99US-0132863.	
PR	11-MAY-1999;	99US-0134256.	
PR	14-MAY-1999;	99US-0134218.	
PR	14-MAY-1999;	99US-0134219.	
PR	14-MAY-1999;	99US-0134221.	
PR	14-MAY-1999;	99US-0134370.	
PR	18-MAY-1999;	99US-0134768.	
PR	19-MAY-1999;	99US-0134941.	
PR	20-MAY-1999;	99US-0135124.	
PR	21-MAY-1999;	99US-0135353.	
PR	24-MAY-1999;	99US-0135629.	

QY 267 TCTTCAAGACGACAGATGTCACATCTCTTTGAAGCGTTGTTGGAACTCTGAAAGCCG 326 99US-0138847.
14-JUN-1999; PR 99US-0139119.
16-JUN-1999; PR 99US-0139452.
16-JUN-1999; PR 99US-0139453.
17-JUN-1999; PR 99US-0139492.
18-JUN-1999; PR 99US-0139454.
18-JUN-1999; PR 99US-0139455.
18-JUN-1999; PR 99US-0139456.
18-JUN-1999; PR 99US-0139457.
18-JUN-1999; PR 99US-0139458.
18-JUN-1999; PR 99US-0139459.
18-JUN-1999; PR 99US-0139460.
18-JUN-1999; PR 99US-0139461.
18-JUN-1999; PR 99US-0139462.
18-JUN-1999; PR 99US-0139463.
18-JUN-1999; PR 99US-0139750.
18-JUN-1999; PR 99US-0139763.
21-JUN-1999; PR 99US-0139817.
22-JUN-1999; PR 99US-0139899.
23-JUN-1999; PR 99US-0140353.
23-JUN-1999; PR 99US-0140354.
24-JUN-1999; PR 99US-0140695.
28-JUN-1999; PR 99US-0140823.
29-JUN-1999; PR 99US-0140991.
30-JUN-1999; PR 99US-0141287.
01-JUL-1999; PR 99US-0141842.
02-JUL-1999; PR 99US-0142154.
02-JUL-1999; PR 99US-0142055.
06-JUL-1999; PR 99US-0142390.
08-JUL-1999; PR 99US-0142803.
09-JUL-1999; PR 99US-0142920.
12-JUL-1999; PR 99US-0142977.
13-JUL-1999; PR 99US-0143542.
14-JUL-1999; PR 99US-0143624.
15-JUL-1999; PR 99US-0144005.
16-JUL-1999; PR 99US-0144085.
16-JUL-1999; PR 99US-0144086.
19-JUL-1999; PR 99US-0144331.
19-JUL-1999; PR 99US-0144332.
19-JUL-1999; PR 99US-0144333.
19-JUL-1999; PR 99US-0144334.
19-JUL-1999; PR 99US-0144335.
20-JUL-1999; PR 99US-0144352.
20-JUL-1999; PR 99US-0144632.
20-JUL-1999; PR 99US-0144884.
21-JUL-1999; PR 99US-0144814.
21-JUL-1999; PR 99US-0145086.
21-JUL-1999; PR 99US-0145088.
22-JUL-1999; PR 99US-0145085.
22-JUL-1999; PR 99US-0145087.
22-JUL-1999; PR 99US-0145089.
22-JUL-1999; PR 99US-0145192.
23-JUL-1999; PR 99US-0145145.
23-JUL-1999; PR 99US-0145218.
23-JUL-1999; PR 99US-0145224.
26-JUL-1999; PR 99US-0145276.
27-JUL-1999; PR 99US-0145913.
27-JUL-1999; PR 99US-0145918.
27-JUL-1999; PR 99US-0145919.
28-JUL-1999; PR 99US-0145951.
02-AUG-1999; PR 99US-0146386.
02-AUG-1999; PR 99US-0146388.
02-AUG-1999; PR 99US-0146389.
03-AUG-1999; PR 99US-0147038.
04-AUG-1999; PR 99US-0147204.
04-AUG-1999; PR 99US-0147302.
05-AUG-1999; PR 99US-0147192.
05-AUG-1999; PR 99US-0147260.
06-AUG-1999; PR 99US-0147303.
07-AUG-1999; PR 99US-0147416.
09-AUG-1999; PR 99US-0147493.
09-AUG-1999; PR 99US-0147935.
DB 270 CTTCTTCGTGATGATAAATGTCACCTCTTTGAAGCATGCTAGGAACCCTTAAGCTG 329
QY 327 CAAACAGGAGGAGATTGTTAC 348
DB 330 CAAACAGGAGGAGATTGTTAC 351
RESULT 14
AAC53910
ID AAC53910 standard; DNA; 551 BP.
XX AC AAC53910;
DT 18-OCT-2000 (first entry)
DE Zea mays DNA fragment SEQ ID NO: 76053.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130445.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135623.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.

Db	260	AAATGCTGATGGAAGTAANTGTGAAATTTGGGGTCTCTTTCTGATGATAAATGAGC	319
QY	289	CAATCTCTTTGAAGCGTTGGTGGGAACCTCTGAAAGCCGCAAAACGAGGAAGATTGT	345
Db	320	CAACCTCTTTGAAGCATTTGGTAGGAACCTCTTAAAGCTGTAAACGAGANANAAGAAATGT	376
RESULT 13			
ABALL1066			
ID	ABALL1066	standard; cDNA; 351 BP.	
XX	XX		
AC	ABALL1066;		
DT	23-JAN-2002	(first entry)	
XX	XX		
DE	Human nervous system related polynucleotide SEQ ID NO 73.		
XX	XX		
KW	Human; nootropic; neuroprotective; cytostatic; dermatological; virucide		
KW	immunosuppressive; antinflammatory; anti-HIV; antitubercular; vulnerary		
KW	antiparkinsonian; antischizoid; antianemic; antidiabetic; cancer;		
KW	antirheumatic; hepatoprotective; cerebroprotective; antiinflammatory;		
KW	antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;		
KW	antiparasitic; cardiant; immune disorder; cardiovascular disorder;		
KW	neurological disease; infection; nephrotropic; gene therapy; vaccine; s		
OS	Homo sapiens.		
XX	XX		
PN	WO200159063-A2.		
XX	XX		
PD	16-AUG-2001.		
XX	XX		
PF	17-JAN-2001; 2001WO-US01334.		
XX	XX		
PR	31-JAN-2000; 2000US-0179065.		
PR	04-FEB-2000; 2000US-0180628.		
PR	24-FEB-2000; 2000US-0184664.		
PR	02-MAR-2000; 2000US-0186350.		
PR	16-MAR-2000; 2000US-0189874.		
PR	17-MAR-2000; 2000US-0190076.		
PR	18-APR-2000; 2000US-0198123.		
PR	19-MAY-2000; 2000US-0205515.		
PR	07-JUN-2000; 2000US-0209467.		
PR	28-JUN-2000; 2000US-0214886.		
PR	30-JUN-2000; 2000US-0215135.		
PR	07-JUL-2000; 2000US-0216647.		
PR	07-JUL-2000; 2000US-0216880.		
PR	11-JUL-2000; 2000US-0217487.		
PR	11-JUL-2000; 2000US-0217496.		
PR	14-JUL-2000; 2000US-0218290.		
PR	26-JUL-2000; 2000US-0220963.		
PR	26-JUL-2000; 2000US-0220964.		
PR	14-AUG-2000; 2000US-0224518.		
PR	14-AUG-2000; 2000US-0224519.		
PR	14-AUG-2000; 2000US-0225213.		
PR	14-AUG-2000; 2000US-0225214.		
PR	14-AUG-2000; 2000US-0225266.		
PR	14-AUG-2000; 2000US-0225267.		
PR	14-AUG-2000; 2000US-0225268.		
PR	14-AUG-2000; 2000US-0225270.		
PR	14-AUG-2000; 2000US-0225447.		
PR	14-AUG-2000; 2000US-0225757.		
PR	14-AUG-2000; 2000US-0225758.		
PR	14-AUG-2000; 2000US-0225759.		
PR	18-AUG-2000; 2000US-0226279.		
PR	22-AUG-2000; 2000US-0226681.		
PR	22-AUG-2000; 2000US-0226688.		
PR	22-AUG-2000; 2000US-0227182.		
PR	23-AUG-2000; 2000US-0227009.		
PR	30-AUG-2000; 2000US-0228924.		
PR	01-SEP-2000; 2000US-0229287.		
PR	01-SEP-2000; 2000US-0229343.		
PR	01-SEP-2000; 2000US-0229344.		
PR	01-SEP-2000; 2000US-0229345.		

Db	260	AAATGCTGATGGAAGTAANTGTGAAATTTGGGGTCTCTTTCTGATGATAAATGAGC	319
QY	289	CAATCTCTTTGAAGCGTTGGTGGGAACCTCTGAAAGCCGCAAAACGAGGAAGATTGT	345
Db	320	CAACCTCTTTGAAGCATTTGGTAGGAACCTCTTAAAGCTGTAAACGAGANANAAGAAATGT	376
RESULT 13			
ABALL1066			
ID	ABALL1066	standard; cDNA; 351 BP.	
XX	XX		
AC	ABALL1066;		
DT	23-JAN-2002	(first entry)	
XX	XX		
DE	Human nervous system related polynucleotide SEQ ID NO 73.		
XX	XX		
KW	Human; nootropic; neuroprotective; cytostatic; dermatological; virucide		
KW	immunosuppressive; antinflammatory; anti-HIV; antitubercular; vulnerary		
KW	antiparkinsonian; antischizoid; antianemic; antidiabetic; cancer;		
KW	antirheumatic; hepatoprotective; cerebroprotective; antiinflammatory;		
KW	antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;		
KW	antiparasitic; cardiant; immune disorder; cardiovascular disorder;		
KW	neurological disease; infection; nephrotropic; gene therapy; vaccine; s		
OS	Homo sapiens.		
XX	XX		
PN	WO200159063-A2.		
XX	XX		
PD	16-AUG-2001.		
XX	XX		
PF	17-JAN-2001; 2001WO-US01334.		
XX	XX		
PR	31-JAN-2000; 2000US-0179065.		

```

Db 193 TCAGGAATCTGGTAGGAATTGCAGCAATGAGATTTTTCGGGGCAAGCATGGGAATG 134
QY 663 TGTTTCTTAAACCAATCTTAGTCT--CTGCCACACTTGACACTCCGTCGAAGTGAGAA 720
Db 133 TTGTTCATAAATAATAGACATTTCTATAGATATTGACATTTCTCGGAAGCAACAAG 74
QY 721 CGAACTAAAGACCAACTCGCGTGGGAAATATTATGTTATGTAATAAAAAA 773
Db 73 CAACATGAAGACCAACTCTCTATGAGAAATATTATGATGTTTATGTAATAAGA 21

RESULT 10
AAZ21551
ID AAZ21551 standard; DNA; 177 BP.
XX
AC AAZ21551;
XX
DT 01-DEC-1999 (first entry)
XX
DE Fragment of progression suppressed gene 13 (PSGen13).
XX
KW Progression suppressed gene; PSGen; progression elevated gene; PSGen;
KW tumour; reciprocal subtraction differential RNA display; RSDD;
KW differential expression; gene cloning; cancer; ss.
XX
OS Rattus sp.
XX
PN WO9943844-Al.
XX
PD 02-SEP-1999.
XX
PE 26-FEB-1999; 99WO-US04323.
XX
PR 27-FEB-1998; 98US-0032684.
XX
PR 03-NOV-1998; 98US-0185115.
XX
PR 23-NOV-1998; 98US-0197889.
XX
PA (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI Fisher PB;
XX
WPI; 1999-550872/46.
XX
DR Identifying nucleic acids differentially expressed between two samples,
PT particularly sequences involved in tumour progression
XX
PS Examples; Fig 23; 110pp; English.
XX
CC This sequence is a fragment of the progression suppressed gene 13
CC (PSGen13). This gene has suppressed expression in progressed tumour
CC cells. PSGen13 was identified using new methods for identifying nucleic
CC acids differentially expressed between two samples. The method involves
CC performing reciprocal subtraction differential RNA display (RSDD) between
CC the two samples to generate two subtraction samples. The subtraction
CC samples are amplified and compared to identify those nucleic acids that
CC are differentially expressed. The method is used to identify and clone
CC differentially expressed genes, particularly those with increased or
CC reduced expression during tumour cell progression, e.g. progression
CC suppressed genes (PSGen) and progression elevated genes (PEGen). The
CC method reduces the complexity of the band pattern produced in
CC conventional differential RNA display (where bands may be obscured,
CC resulting in false positive signals) since most bands common to both
CC samples are eliminated, allowing identification and cloning of genes
CC displaying anticipated differential expression. RSDD requires only a
CC single anchored primer for amplification and reamplified cDNA can be
CC analysed by reverse Northern blotting.
XX
SQ Sequence 177 BP; 72 A; 28 C; 30 G; 47 T; 0 other;
Query Match 20.1%; Score 157; DB 20; Length 177;
Best Local Similarity 100.0%; Pred. No. 2.5e-34;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 624 GTAGGCAATAAATCTTTTCAGAGGTGCGAAAAAGCTTTTGTCTTAAACCATCTTA 683
Db 1 GTAGGCAATAAATCTTTTCAGAGGTGCGAAAAAGCTTTTGTCTTAAACCATCTTA 60
QY 684 GTCCTCTGCCACACTTGACACTCCGTCGAAGTGAGAGCGAACTAAAGACCAACTCGG 743
Db 61 GTCCTCTGCCACACTTGACACTCCGTCGAAGTGAGAGCGAACTAAAGACCAACTCGG 120
QY 744 GAAATATTATGTTTATGTAATAAAAAAATCATGT 780
Db 121 GAAATATTATGTTTATGTAATAAAAAAATCATGT 157

RESULT 11
AAZ25869
ID AAZ25869 standard; cDNA; 177 BP.
XX
AC AAZ25869;
XX
DT 08-JUN-1999 (first entry)
XX
DE Clone PSGen-13 sequence.
XX
KW Isolation; identification; reciprocal subtraction differential display;
KW RSDD; gene expression; ss.
XX
OS Mus sp.
XX
PN US5882874-A.
XX
PD 16-MAR-1999.
XX
PE 27-FEB-1998; 98US-0032684.
XX
PR 27-FEB-1998; 98US-0032684.
XX
PA (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI Fisher PB;
XX
WPI; 1999-214060/18.
XX
DR Identifying differentially expressed nucleic acids between two
PT samples - by reciprocal subtraction
XX
PS Examples; Fig 23; 29pp; English.
XX
CC The invention relates to a method of isolating and identifying
CC differentially expressed nucleic acids between two samples by using
CC a reciprocal subtraction differential display (RSDD) method. This
CC sequence represents clone PSGen-13 which is a novel gene.
XX
SQ Sequence 177 BP; 72 A; 28 C; 30 G; 47 T; 0 other;
Query Match 20.1%; Score 157; DB 20; Length 177;
Best Local Similarity 100.0%; Pred. No. 2.5e-34;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 624 GTAGGCAATAAATCTTTTCAGAGGTGCGAAAAAGCTTTTGTCTTAAACCATCTTA 683
Db 1 GTAGGCAATAAATCTTTTCAGAGGTGCGAAAAAGCTTTTGTCTTAAACCATCTTA 60
QY 684 GTCCTCTGCCACACTTGACACTCCGTCGAAGTGAGAGCGAACTAAAGACCAACTCGG 743
Db 61 GTCCTCTGCCACACTTGACACTCCGTCGAAGTGAGAGCGAACTAAAGACCAACTCGG 120
QY 744 GAAATATTATGTTTATGTAATAAAAAAATCATGT 780
Db 121 GAAATATTATGTTTATGTAATAAAAAAATCATGT 157

RESULT 12

```

PD 11-APR-2002.
 XX
 PF 02-OCT-2001; 2001WO-US30732.
 XX
 PR 02-OCT-2000; 2000US-237271P.
 XX
 PA (FARB) BAYER CORP.
 XX
 XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
 PI Thlaglingam A, Lewis ME;
 PI WPI; 2002-426115/45.
 XX
 DR New isolated nucleic acid that is differentially expressed in cancer
 XX tissues useful for determining the presence of colon cancer in a cell
 XX or tissue type, and in antisense therapy
 XX
 XX Claim 1; Fig 1; 796pp; English.
 XX
 CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
 CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
 CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
 CC used in antisense therapy. An antibody immunoreactive with a polypeptide
 CC encoded by (I) is useful for detecting cancer in a patient sample, and
 CC for detecting the presence or absence of a polynucleotide encoded by a
 CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
 CC from (I) can be used for determining the presence of a nucleic acid which
 CC hybridises to (I), and for determining the phenotype of cells in a sample
 CC of cells from a patient. (I) is useful for determining the presence of
 CC colon cancer in a cell or tissue type, for determining the presence or
 CC state of other type of cancer, in antisense therapy, to generate
 CC macroarrays on a solid surface, to identify a chromosome on which the
 CC corresponding gene resides, and in tissue profiling, forensics, genetic
 CC analysis, mapping and diagnostic applications. (I) can be used to raise
 CC antibodies, and to screen for peptide analogues and antagonists.
 XX
 XX Sequence 627 BP; 171 A; 151 C; 153 G; 147 T; 5 other;
 SQ
 Query Match 27.6%; Score 215.6; DB 24; Length 627;
 Best Local Similarity 83.3%; Pred. No. 1.3e-50;
 Matches 245; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 QY 147 AGGAAGCAATCGGAGGTGACGATGACGAGGAGGATGAGTAACTCTCTGGTGGAGG 206
 Db AAGAGGAGCTACGATGCGAGCAATGAATGGATCAGGAGTTAACTCTTAGTGGAGG 235
 QY 207 AAATTCATCGTCTGGTTCCTCAAAATGCGATGGAACCTGAGTGTGAAGTTGGGTC 266
 Db AAATTCATCGTCTGGTTCCTCAAAATGCGATGGAACCTGAGTGTGAAGTTGGGTC 175
 QY 267 TCTTCCAGACGACAGATGTCCTATCTTTGAAGCGTTGGTGGAACTCTGAAAGCG 326
 Db TCTTCCGATGATGAATGTGCAACCTCTTTGAAGCATTGGTGAAGTCTTAAAGCTG 115
 QY 327 CAACACGAAGCAAGATTCTTACGACGAGAGAGCTGCTTTGCAAGGTTTCATGATG 386
 Db CAACACGAAGCAAGATTCTTACATATACAGGAGAGCTGCTTTCGAAGGTTTCATGATG 55
 QY 387 ATGTTGACATGTTATGTCGCAAGATTAAATGTTGTTTCAGATCTGGGGTATC 440
 Db ATGTTGACATTAATATGTTGCAAGATTAAATGTTGTTTACATATCTTATGATACC 1
 RESULT 9
 ABK45388/C
 ID ABK45388 standard; cDNA; 552 BP.
 XX
 AC ABK45388;
 XX
 XX 05-JUN-2002 (first entry)
 DT
 XX cDNA encoding colon tumour protein, SEQ ID No 939.
 DE
 XX

KW Human; colon tumour; vaccine; colon cancer; immunogenic;
 KW immunotherapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200212328-A2.
 XX
 PD 14-FEB-2002.
 XX
 PF 31-JUL-2001; 2001WO-US24218.
 XX
 PR 03-AUG-2000; 2000US-223283P.
 PR 28-MAR-2001; 2001US-279763P.
 PR 29-JUN-2001; 2001US-302051P.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 XX King GE, Meagher MJ, Xu J, Secrlist H;
 PI WPI; 2002-241739/29.
 DR
 XX New colon cancer polypeptides and polynucleotides, useful as vaccines,
 XX for diagnosing, preventing, and treating colon cancer, and as markers
 XX for the progression of cancer -
 XX
 XX Claim 1; SEQ ID No 939; 147pp; English.
 PS
 CC The invention relates to polynucleotides encoding colon tumour proteins.
 CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
 CC compositions, such as vaccines, for the diagnosis, prevention, and
 CC treatment of colon cancer. Polynucleotide sequences may be used as
 CC hybridisation probes or primers, and in the design and preparation of
 CC ribozyme molecules for inhibiting expression of tumour polypeptides and
 CC proteins in tumour cells. The compositions are useful for stimulating an
 CC immune response against cancer, particularly for the immunotherapy of
 CC colon cancer, and as markers for the progression of cancer.
 CC ABK44450-ABK46237 represent coding sequences of human colon tumour
 CC proteins of the invention.
 CC Note: With the exception of SEQ ID No 1 and 2, the sequence data
 CC for this patent did not form part of the printed specification but was
 CC supplied by the European Patent Office.
 XX
 XX Sequence 552 BP; 190 A; 104 C; 75 G; 183 T; 0 other;
 SQ
 Query Match 26.5%; Score 206.6; DB 24; Length 552;
 Best Local Similarity 68.7%; Pred. No. 4.4e-48;
 Matches 366; Conservative 0; Mismatches 144; Indels 23; Gaps 5;
 QY 263 GTCTCTTCCAGACGACAGATGTCCTTTGAAGCGTTGGTGGAACTCTGAAA 322
 Db GTCTCTTCCGATGATGATAAATGTGCAACCTCTTTGAAGCATTGGTAGAACTCTTAAA 493
 QY 323 GCGGCAAAAGGAGGAGATGTTAGCTAGCAGGAGAGCTGCTTTTGCAGGTGTTCA 382
 Db GCTGCAAAAGGAGGAGATGTTAGCATATCCAGAGAGCTGCTTCTGCAAGGTGTTCA 433
 QY 383 GATGATGTTGACATGTTATGCTGCAAGATTAAATGTTGCAAGTCTGGGGTA---- 438
 Db GATGATGTTGACATGTTATGTTGCAAGATTAAATGTTGCAAGTCTGGGGTA---- 373
 QY 439 -----TCTGTAACCTGGAATAAATTAAGTTAAAGGACAAACAT---GAAGTTC 484
 Db CATTTTTTGTCTGTAACCTGGAATAAATTAAGTTAAAGGACAAACATTTGAACATACT 314
 QY 485 TTATGTTATTTTATACCTTTGTAACAAAGGGGA--CTGTTGAGAGTCTCTGTTTT 542
 Db TAATGTTATTTTATAGAACTTTGTAACCAAGGAGGATTCATGTTTATAGAGCTCTGCT 254
 QY 543 TATACCTTGGAGCAAAACATTTACATGTAATAAATAACAAACCTGTTATTTTTC 602
 Db TTTTATATCTTGAAAGAAATCTATGATGATGATGATGATGATGATGATGATGATGATG 194
 QY 603 TTAAGAAGGTAATCGGGAGAGCTAGGCAATAAATGTTTTTTCAGAGGTGCGAAAAAGCTT 662

Query Match 32.3%; Score 252.2; DB 21; Length 712;
Best Local Similarity 76.1%; Pred. No. 6.8e-61;
Matches 357; Conservative 0; Mismatches 93; Indels 19; Gaps 3;

QY 73 GCTTGTTCCTCCCTAGGCGCGGAAGCTGAGTGCAGGTTTCAGACCCACCGCGGCAGCAG 132
DB 32 GCCTCACCCTCCCGAGTGCAGTGAAGAGTAAACCGGTTCAGACCCACCGCGGCAGT 91
QY 133 CTCTTCAGTGAAGAAGCAATCGGAGGTCAGCAATGAACGTGAGCATGAGGTTAA 192
DB 92 TCTCCGCGGGAAGAAACCGCGCAGAGGAGCAATGAATGTGATCAGAGGTTAA 151
QY 193 CTTCTGCTGGTGAAGAAATTCATGCTCTGGGTTCCAAAATGCCGATGGGAAACTGAGTGT 252
DB 152 CTTCTAGTGGAGAAATTCATGCTCTGGGTTCCAAAATGCTCATGGAAGTTAAGCGT 211
QY 253 GAAGTTTGGGGTCTCTTCCAAAGACGACATGTGCCAATCTCTTTGAAGGTTGGTGGG 312
DB 212 GAATTTGGGGTCTCTTCCGATGATGAATGTGCCAATCTCTTTGAAGCATTTGGTAGG 271
QY 313 AACTCTGAAGCGCGCAAAAGAAAGATTTGTTACCTAGCAGGAGGCTGCTTTTGCA 372
DB 272 AACTCTTAAAGCTGCAAAAGCAAGAGATTTGTAACATATCCAGGAGGCTGCTTCTGCA 331
QY 373 AGTGTTTCATGATGTTGACATGTGATGCTGCAAGATTAATGTGTTGCAAGATCTG 432
DB 332 AGTGTTTCATGATGTTGACATTTATTTACTGCAAGATTAATGTGTTGCAATATCTT 391
QY 433 GGGGTA-----TCTGGTAAACTGGAATTAATTAAGTTAAAGGACAAACAT- 476
DB 392 TATGTACTGCCATTTTGTGTTCTGGTAAACTGGAAT-ATAAGTGAAGAACAAACATT 450
QY 477 --GAAGTTCCTTATGATTTTATAGACCTTTGTTAAACAAAAGGGGACT 523
DB 451 TGAACATACCTTAATGATTTTATAGAACCTTTGTTAAACGAAGAGATT 499

RESULT 7
ABQ57053
ID ABQ57053 standard; cdna; 717 BP.
AC ABQ57053;
XX
XX
DT 02-AUG-2002 (first entry)
XX
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:748.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200229086-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30732.
XX
PR 02-OCT-2000; 2000US-237271P.
XX
PA (FARB) BAYER CORP.
XX
PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI Thiagalingam A, Lewis ME;
XX
DR WPI; 2002-426115/45.
XX
PT New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell
PT or tissue type, and in antisense therapy
XX
PS Claim 1; Fig 1; 796pp; English.
XX

ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridizes to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridizes to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence of
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists.
XX
SQ Sequence 717 BP; 192 A; 171 C; 155 G; 180 T; 19 other;

Query Match 32.1%; Score 250.4; DB 24; Length 717;
Best Local Similarity 75.6%; Pred. No. 2.2e-60;
Matches 323; Conservative 0; Mismatches 102; Indels 2; Gaps 1;

QY 16 GGTCCCTCCCTTCTCCACTGCAGCCTTCTCTTAGCCGGAACCACTTCTCTTCTCTGCT 75
DB 48 CTTGCTCTCTCTCTCCACAGCGCTTCTCTTTGCTTCAGCCACTTCTCTTCTTGGCC 107
QY 76 TGTTCCTCCTTAGGCGCGGAAGCTGAGTGCAGGTTTCAGCCCGCGGAGCAGCTC 135
DB 108 TCACCCCTCCCGAGTCACTGAAGAAGGTAAACCGGGTCCAGACCCACCGCGCCAGTTCT 167
QY 136 TTCAGTGAAGAAGGAAGCAATCGGAGGTCAGCAATGAACGTGGAGCATGAGTTAACTT 195
DB 168 CCGCGGGAAGGAAGAAACCGCGCAGAGAGCAGCAATGAATGTGGATCAGAGGTTAACTT 227
QY 196 CTTGTTGGAGGAATTCATGCTGCTGGGTTCCAAAATGCCGATGGGAACTGAGTGTGAA 255
DB 228 CTTAGTGGAGGAATTCATGCTGTTGGGTTTCAAAAATGCTGATGGAAGTTAAGCGTGA 287
QY 256 GTTGGGGTCTCTTCCAAAGCAGCAGATGTGCCAATCTCTTTGAAGCGTTGGTGGGAAC 315
DB 288 ATTTGGGGTCTCTTCCGATGATGATAAATGTGCCAACCTCTTTGAAGCATTTGTAGGAAC 347
QY 316 TCTGAAGCCGCAAAACGAAGGAGATTGTTACGTACGAGGAGGAGCTCTTTTTCGAAGG 375
DB 348 TCTTAAACCTGAAACCAAGGAAGAGATTGTAACATATCCAGGAGAGCTCTTCTCAAGG 407
QY 376 TGTTCATGATGATGTTGACATTTGATGCTGCAA--GATTAATGTGGTTGCGAGATCTGG 433
DB 408 TGTTCATGATGATGTTGACATTTATATTACTGCAAGATAAATGNGGGTTACATATCTTT 467
QY 434 GGGTATC 440
DB 468 ATGTACC 474

RESULT 8
ABQ56862/c
ID ABQ56862 standard; cdna; 627 BP.
XX
XX
AC ABQ56862;
XX
XX
DT 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:557.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200229086-A2.
XX

[illegible]

Db 343 TCCTTAAGCTGCAAAACGAAGAGATGTAACATATCCAGGAGAGCTGCTTCTGCAAGG 402
 QY 376 TGTTTCATGATGTTGACATTTGATTCGCAAGATTAATGTTGCGAGATCTGGGG 435
 Db 403 TGTTTCATGATGTTGACATTTGATTCGCAAGATTAATGTTGCGAGATCTGTTAT 462
 QY 436 GTA-----TCTGTAACCTGGAATTAATGTTGCAAGATTAATGTTGCGAGATCTGTTAT 477
 Db 463 GTACTGCCATTTTGTGTTCTGTAACCTGGAAT-ATAAAGTGAAGAACAACAATTTGA 521
 QY 478 AAGTCTCTATGATTTTATAGACCTTTGTAACAAAGGGGACT--TCTTGAGAGTC 535
 Db 522 ACATCTTAATGATTTTATAGACCTTTGTAACAAAGGGGAGATTCATGTTTAGAAGT 581
 QY 536 CTGTTTTTATACCTTGAGCAAAACATTAACAATGTAATAAATAACAACCTGTTATTT 595
 Db 582 CTGCTCTTTTATATCTTGAAAGAAATCTATGATGATGCTATATAAATAAATCTTAT 641
 QY 596 TTTTCTTAAGAGGTAATCGGACAGCTGAGCAATAAATGTTTTCAGAGGTCGAAA 655
 Db 642 ATTTTCTCAGGAATCTGTTAGGAATTCGAGCAATGAGATTTTTCGCGGGCAGGAT 701
 QY 656 AAGCTTTTGTGTTTCTTAACCATCTTAGTCT--CTGCCACACTTGACACTCGTCAAG 713
 Db 702 GGAATGTTTGTCAATAAATAATAGACATTTTCTATAGATATTTGACATCTCGGAAG 761
 QY 714 TGAGAAGCAACTTAAGACCAACTCGGGTGGAAATATATGTTTATGTAATAAATAA 773
 Db 762 CAACAAGCAACTGAAGACCAACTCTATGAGAAATATATGATGTTTATGTAATAAGA 821

RESULT 4

AAFI6110
 ID AAFI6110 standard; cdna; 778 BP.
 AC AAFI6110;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:545.
 XX
 KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
 KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KW wound; infectious disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO20005174-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05988.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2000-587513/55.
 DR P-PSDB; AAB56907.
 XX
 PT Prostate cancer associated gene sequences, referred to as prostate
 PT cancer antigens, useful for treatment, prevention, and diagnosis of
 PT disorders such as prostate cancer -
 XX
 PS Claim 1; Page 1004-1005; 2338pp; English.
 XX
 CC AAFI5566 to AAFI6505 encode the human prostate cancer associated

CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
 CC The prostate cancer antigens can have neuroprotective, cytostatic,
 CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
 CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen
 CC polynucleotides may be used for detection of prostate cancer, chromosome
 CC identification, as chromosome markers, and for numerous other, diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAFI6506 to AAFI6514 to
 CC AAB57303 represent sequences used in the exemplification of the present
 CC invention.

XX
 SQ Sequence 778 BP; 250 A; 151 C; 148 G; 221 T; 8 other;

Query Match 36.9%; Score 288.2; DB 21; Length 778;
 Best Local Similarity 69.9%; Pred. No. 5e-71;
 Matches 451; Conservative 2; Mismatches 167; Indels 25; Gaps 4;

QY 16 GGTCCCTCCCTTCCCACTGCAGCTTCTTAGCCCGAACCACTTCTTCTTCTGCT 75
 Db 32 CCTGCTCTCTCTCCACAGCGTCTTCTTGTGCTCAGCCACTTCTTCTTCTGCCC 91
 QY 76 TGTTCCTCCTTAGGCGCGGAGCTGAGTGCAGGGTTTCCAGCCCGGCGGAGCAGCTC 135
 Db 92 TCACCTCCCGCTGCTGCTGAGAGGTAACCGGCTCAGACCCCGGCGGCGGAGTCT 151
 QY 136 TTAGTGAAGAGGAGCAATCGGAGGCTGAGCAATGAACGTTGGAGCATGAGTTAACT 195
 Db 152 CCGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 211
 QY 196 CCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 255
 Db 212 CTTAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 271
 QY 256 GTTGGGGTCTCTTCCAAAGACAGAGATGTCCTTGAAGCGTTGGTGGGAGAC 315
 Db 272 ATTTGGGGTCTCTTCCGATGATGATGATGATGATGATGATGATGATGATGATG 331
 QY 316 TCTGAAGCCCAAAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 375
 Db 332 TCTTAAAGCTGCAAAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 391
 QY 376 TGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 435
 Db 392 TGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 451
 QY 436 GTA-----TCTGTAACCTGGAATTAATGTTAAGGAGCAAAACAT---G 477
 Db 452 GTACTGCCATTTTGTGTTCTGTTAACTGGAAT-ATAAAGTGAAGAACAACAATTTGA 510
 QY 478 AAGTCTCTATGATTTTATAGACCTTTGTAACAAAGGGGACT--TCTTGAGAGTC 531
 Db 511 ACATCTTAATGATTTTATAGACCTTTGTAACAAAGGGGAGATTCATGTTTAGAAGT 570
 QY 532 AGTCTGTTTTTATACCTTGAGCAAAACATTAACAATGTAATAAATAACAACCTGTTA 591
 Db 571 CTGCTCTTTTATATCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 630
 QY 592 TTTTCTTTTCTTAAAGAGGTAATCTCGGAGAGCTAGGCAATAAATAA 636
 Db 631 ATTTCTTGAAGAGGTAATCTCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 675

RESULT 5

ABQ57055
 ID ABQ57055 standard; cdna; 594 BP.
 XX
 AC ABQ57055;
 XX
 DT 02-AUG-2002 (first entry)
 XX

Query Match	96.8%	Score 754.8	DB 20	Length 800
Best Local Similarity	99.5%	Pred. No. 1.3e-202		
Matches 778	Conservative 0	Mismatches 2	Indels 2	Gaps 2
Qy 1	GGCAGGAGCTCTCCCTCGTCCCTCCCTCTCTCCACTGCAGCCTTTCTCTTACGCCGAACCA	60		
Db 1	GGCAGGAGCTCTCCCTCGTCCCTCCCTCTCTCCACTGCAGCCTTTCTCTTACGCCGAACCA	60		
Qy 61	CTTCCTCTTCTGCTGTGTTCTCCTTAGGGCGCGGAAGCTGAGTGCAGGGTTCAGACCCA	120		
Db 61	CTTCCTCTTCTGCTGTGTTCTCCTTAGGGCGCGGAAGCTGAGTGCAGGGTTCAGACCCA	120		
Qy 121	CGCGGCGAGCAGCTCTCAGTGAAGAAGGAAGCAATCGGAGGGTCAGCAATGAACGTGGA	180		
Db 121	CGCGGCGAGCAGCTCTCAGTGAAGAAGGAAGCAATCGGAGGGTCAGCAATGAACGTGGA	180		
Qy 181	GCATGAGGTTAACTCCCTGGTGGAGGAAATCATCTGCTGGGTTCGCAAAATGCCGATGG	240		
Db 181	GCATGAGGTTAACTCCCTGGTGGAGGAAATCATCTGCTGGGTTCGCAAAATGCCGATGG	240		
Qy 241	GAAACTGAGTGTGAAGTTTGGGGTCTCTTCCAAGCAGACAGATGTGCCAATCTCTTTGA	300		
Db 241	GAAACTGAGTGTGAAGTTTGGGGTCTCTTCCAAGCAGACAGATGTGCCAATCTCTTTGA	300		
Qy 300	AAGCGTGGTGGGAACCTCTGAAAG-CGCGAAACGAGGAAGATGTTACGTACGCAGGA	358		
Db 301	AACGTTGGTGGGAACCTCTGAAAGCCGCAAAAGGAAGATGTTACGTACGCAGGA	360		
Qy 359	GAGCTGCTTTTGGCAAGGTGTTTCATGATCATCTTGACATTTGATGCTGCAAGATTAATGT	418		
Db 361	GAGCTGCTTTTGGCAAGGTGTTTCATGATCATCTTGACATTTGATGCTGCAAGATTAATGT	420		
Qy 419	GGTTGCGAGATCTGGGGGTATCTGGTAACTGGAATTAATTAAGTTAAAGSACAACATGA	478		
Db 421	GGTTGCGAGATCTGGGGGTATCTGGTAACTGGAATTAATTAAGTTAAAGSACAACATGA	480		
Qy 479	AGTTCCCTTATCTATTTTATAGCCTTGTAAACAAAGGGGACTGTTGTGAGAAGTCTTG	538		
Db 481	AGTTCCCTTATCTATTTTATAGCCTTGTAAACAAAGGGGACTGTTGTGAGAAGTCTTG	540		
Qy 539	TTTTTATACCTTGGAGCAAAACATTACAATGTAATAAATAAACAACCTGTTATTTTTTT	598		
Db 541	TTTTTATACCTTGGAGCAAAACATTACAATGTAATAAATAAACAACCTGTTATTTTTT	600		
Qy 599	TTTCTTAAGAAGGTAATCGGAGAGCTAGGCAATAAATGTTTTCAGAGGTGCGAAAAAG	658		
Db 601	TTTCTTAAGAAGGTAATCGGAGAGCTAGGCAATAAATGTTTTCAGAGGTGCGAAAAAG	660		
Qy 659	CTTTTGTGTTTCTTAAACCATCTTGTAGTCTCTGCGACACTTCGACACTCCGTCAAAAGTGAGA	718		
Db 661	CTTTTGTGTTTCTTAAACCATCTTGTAGTCTCTGCGACACTTCGACACTCCGTCAAAAGTGAGA	720		
Qy 719	AGCGAATTAAGACCAACTGCGGTGGAATAATATGTTATGTAATAAAAAAAATCAT	778		
Db 721	AGCGAATTAAGACCAACTGCGGTGGAATAATATGTTATGTAATAAAAAAAATCAT	780		
Qy 779	GT 780			
Db 781	GT 782			
RESULT 3				
ABK11086				
ID	ABK11086 standard; cdNA, 835 BP.			
AC	ABK11086;			
XX				
DT	05-JUN-2002 (first entry)			
XX				
DE	Human Progression Suppressed Gene 13 (HuPGen 13), cdNA.			
XX				
KW	Human; Progression Suppressed Gene 13; rPGen 13; cytostatic; cancer.			

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 17:20:02 ; Search time 302 Seconds
(without alignments)
5816.419 Million cell updates/sec

Title: US-09-648-310-1
Perfect score: 780
Sequence: 1 ggcacgagctctctctctcc.....gtataaaaaaaatcatgt 780

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	780	100.0	780	24 ABK11085	Rat Progression Suppressed Gene 13
2	754.8	96.8	800	20 AA221517	Progression suppressor
3	335.2	43.0	835	24 ABK11086	Human Progression
4	288.2	36.9	778	21 AAF16110	Human prostate cancer
5	263	33.7	594	24 ABQ57055	Human colon cancer
6	252.2	32.3	712	21 AAC66413	Human secreted protein
7	250.4	32.1	717	24 ABQ57053	Human colon cancer
8	215.6	27.6	627	24 ABQ56862	Human colon cancer
9	206.6	26.5	552	24 ABK45388	cDNA encoding colo

10	157	20.1	177	20	AA221551	Fragment of progne
11	157	20.1	177	20	AA225869	Clone pSGen-13 seq
12	122.4	15.7	388	21	AAC98308	Human colon cancer
13	119.6	15.3	351	22	ABA11066	Human nervous syst
14	114.6	14.7	551	21	AAC53910	zebra mays DNA fragm
15	109.6	14.1	657	21	AAC53673	Arabidopsis thalia
16	92.8	11.9	368	24	ABL38512	Human colon tumour
17	92.2	11.8	406	24	ABN94189	Gene #687 used to
18	87.4	11.2	190	20	AA85712	Novel cDNA sequenc
19	87.4	11.2	190	20	AA85712	Novel cDNA sequenc
20	87.4	11.2	268	24	ABL72646	Corn tassal-derive
21	63.2	8.1	289	24	ABL74776	Corn tassal-derive
22	61.8	7.9	296	21	ABC30197	Human secreted pro
23	46.4	5.9	15161	24	ABL70457	Chemically treated
24	46.4	5.9	15161	24	AA561422	Human gene regulat
25	45	5.8	6956	24	ABL70226	Chemically treated
26	43.4	5.6	11155	24	ABL32605	Human immune syste
27	43.2	5.5	34688	24	ABQ67059	Human angio genesis
28	42.2	5.4	18183	22	AA546279	Tumour suppressor
29	42.2	5.4	18183	24	ABL70111	Chemically treated
30	42.2	5.4	18183	24	ABK31158	Signal transductio
31	41.2	5.3	1008	24	ABN74038	Bovine embryonic g
32	41	5.3	7589	24	ABK28389	DNA transcription
33	40.8	5.2	60	24	ABN49354	Human spliced tran
34	40.6	5.2	5371	22	AA546799	Tumour suppressor
35	40.6	5.2	61020	22	AA546787	Tumour suppressor
36	40.2	5.2	21537	24	ABL33998	Human immune syste
37	40	5.1	611	21	AAA16371	Human colon cancer
38	40	5.1	6452	24	ABN80274	Human chemically m
39	39.8	5.1	16228	24	ABL70460	Chemically treated
40	39.8	5.1	16228	24	AA561425	Human gene regulat
41	39.6	5.1	5371	24	ABL33186	Human immune syste
42	39.6	5.1	5945	24	ABL32084	Human immune syste
43	39.6	5.1	24939	24	ABL70569	Chemically treated
44	39.4	5.1	641	22	ABL19275	Human breast canc
45	39.4	5.1	7104	24	ABL70396	Chemically treated

ALIGNMENTS

RESULT 1	
ABK11085	
ID	ABK11085 standard; cDNA; 780 BP.
XX	
AC	ABK11085;
XX	
DT	05-JUN-2002 (first entry)
XX	
DE	Rat Progression Suppressed Gene 13 (rPSGen 13), cDNA.
XX	
KW	Rat; Progression Suppressed Gene 13; rPSGen 13; cytostatic; cancer;
KW	blood vessel; nasopharyngeal tumour; thyroid tumour; leukaemia;
KW	lymphoma; breast; lung; prostate; ovary; colon; gene; ss.
XX	
OS	Rattus sp.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	170..415
FT	/*tag= a
FT	/product= "progression suppressed gene 13 protein"
XX	
PN	WO200216419-A2.
XX	
PD	28-FEB-2002.
XX	
PF	27-AUG-2001; 2001WO-US26795.
XX	
PR	25-AUG-2000; 2000US-0648310.
XX	
PA	(UYCO) UNIV COLUMBIA NEW YORK.
XX	
PI	Fisher PB, Kang D, Su Z;

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)

DR EMBL: U13194; AAA68468.1; --
 DR EMBL: X83413; CAA58369.1; --
 DR InterPro: IPR003840; Herpes_Helicase.
 DR Pfam: PF02689; Herpes_Helicase; 1.
 KW DNA replication; ATP-binding; Helicase.
 FT NP_BIND 90 97 ATP (BY SIMILARITY).
 SQ SEQUENCE 824 AA; 93286 MW; 5A5507544E06184C CRC64;

Query Match 14.8%; Score 60; DB 1; Length 824;
 Best Local Similarity 23.6%; Pred. No. 40;
 Matches 21; Conservative 14; Mismatches 30; Indels 24; Gaps 3;

QY 1 MNVEHEVNLVVEEHRIGSKNADKGL-----SVKFGVLFQDDRCANL----- 42
 Db 19 MSSAPKTELIVDKVASLSERLEGLPDPWFRHMDPETEFSEFADALCIGIDEFAQPL 78
 QY 43 ----FEALV--GTLKAARRKIVTYAGEL 65
 Db 79 FLPLPKALLVTGTAGAGKTNISQIFLAANL 107

RESULT 15
 CALM_YEAST
 ID CALM_YEAST STANDARD; PRT; 147 AA.
 AC P06787;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 15-JUN-1988 (Rel. 06, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Calmodulin (Cam).
 GN CM1 OR YBR109C OR YBR0904.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87028234; PubMed=3533275;
 RA Davis T.N., Urdea M.S., Maslarsz F.R., Thorner J.;
 RT "Isolation of the yeast calmodulin gene: calmodulin is an essential
 protein.";
 RL Cell 47:423-431(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=95208357; PubMed=7900426;
 RA Mannhaupt G., Stucka R., Ehmlé S., Vetter I., Feldmann H.;
 RT "Analysis of a 70 kb region on the right arm of yeast chromosome II.";
 RL Yeast 10:1363-1381(1994).
 RN [3]
 RP SEQUENCE OF 15-30 AND 128-145.
 RX MEDLINE=87228267; PubMed=3295478;
 RA Davis T.N., Thorner J.;
 RT "Isolation of the yeast calmodulin gene using synthetic
 oligonucleotide probes.";
 RL Meth. Enzymol. 139:248-259(1987).
 RN [4]
 RP CALCIUM-BINDING DATA.
 RX MEDLINE=93278279; PubMed=1304352;
 RA Brockerhoff S.E., Edmonds C.G., Davis T.N.;
 RT "Structural analysis of wild-type and mutant yeast calmodulins by
 limited proteolysis and electrospray ionization mass spectrometry.";
 RL Protein Sci. 1:504-516(1992).
 RN [5]
 RP MUTAGENESIS.
 RX MEDLINE=91256315; PubMed=2044154;
 RA Geiser J.R., van Tuinen D., Brockerhoff S.E., Neff M.M., Davis T.N.;
 RT "Can calmodulin function without binding calcium?";
 RL Cell 65:949-959(1991).

[6]

RN STRUCTURE BY NMR OF 1-78.
 RP MEDLINE=96424985; PubMed=8827436;
 RX Ohki S.-Y., Miura K., Saito M., Nakashima K.-I., Maekawa H.,
 RA Yazawa M., Tsuda S., Hikichi K.;
 RT "Secondary structure and Ca(2+)-binding property of the N-terminal
 half domain of calmodulin from yeast *Saccharomyces cerevisiae* as
 studied by NMR.";
 RL J. Biochem. 119:1045-1055(1996).
 CC -|- FUNCTION: CALMODULIN MEDIATES THE CONTROL OF A LARGE NUMBER OF
 CC ENZYMES BY CA(++) . AMONG THE ENZYMES TO BE STIMULATED BY THE
 CC CALMODULIN-CA(++) COMPLEX ARE A NUMBER OF PROTEIN KINASES AND
 CC PHOSPHATASES.
 CC -|- PTM: THE N-TERMINUS IS BLOCKED.
 CC -|- MISCELLANEOUS: THIS PROTEIN HAS THREE FUNCTIONAL CALCIUM-BINDING
 CC SITES.
 CC -|- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)

DR EMBL: M14760; AAA34504.1; --
 DR EMBL: X78993; CAA55612.1; --
 DR EMBL: Z35978; CAA85064.1; --
 DR PIR: A25060; MCBY.
 DR HSSP: P02593; ICDM.
 DR SGD: S0000313; CMD1.
 DR InterPro: IPR002048; EF-hand.
 DR Pfam: PF00036; ehand; 4.
 DR ProDom: PD000012; EF-hand; 2.
 DR SMART: SM00054; EFh; 3.
 DR PROSITE: PS00018; EF_HAND; 3.
 KW Calcium-binding; Repeat.
 FT CA_BIND 21 32 EF-HAND 1.
 FT CA_BIND 57 68 EF-HAND 2.
 FT CA_BIND 94 105 EF-HAND 3.
 FT DOMAIN 129 140 ANCESTRAL CALCIUM SITE 4.
 FT MUTAGEN 21 21 D->A: HIGHLY REDUCED AFFINITY FOR CA++.
 FT MUTAGEN 32 32 E->V: HIGHLY REDUCED AFFINITY FOR CA++.
 FT MUTAGEN 57 57 D->A: HIGHLY REDUCED AFFINITY FOR CA++.
 FT MUTAGEN 68 68 E->V: HIGHLY REDUCED AFFINITY FOR CA++.
 FT MUTAGEN 94 94 D->A: HIGHLY REDUCED AFFINITY FOR CA++.
 FT MUTAGEN 105 105 E->V: HIGHLY REDUCED AFFINITY FOR CA++.
 SQ SEQUENCE 147 AA; 16135 MW; 819ED1AD5D9400D3 CRC64;

Query Match 14.7%; Score 59.5; DB 1; Length 147;
 Best Local Similarity 27.9%; Pred. No. 6.8;
 Matches 24; Conservative 11; Mismatches 30; Indels 21; Gaps 3;

QY 4 EHEVNLLVEEHRIGSKNADGKLSVKFG-----VLFQDDRCANLFEALV-----G 48
 Db 46 EAEVDNLNNEI-----DVDGNHQIEFSFLALMSQLKSNDESEQLLEAFKVFDRKNGDG 99
 QY 49 TLKAARRKIVTYAGELLQGVHDDV 74
 Db 100 LLSAAELKHVLTSGEKLTDAEVDDM 125

Search completed: January 17, 2003, 02:14:26
 Job time : 81 secs

```

[1]
RN SEQUENCE FROM N.A.
RA Porter M.B., Green M.L., Simmen R.C.M., Sharp D.C.;
RT "Cloning and sequencing of the equine pituitary gonadotropin-releasing
RT hormone receptor and expression of its mRNA following administration
RT of pulsatile or continuous GRH.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH
CC G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL; AF018072; AAC27099.2;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHOOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PSS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 38 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 39 59 1 (POTENTIAL).
FT DOMAIN 60 84 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 85 105 2 (POTENTIAL).
FT DOMAIN 106 115 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 116 136 3 (POTENTIAL).
FT DOMAIN 137 157 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 158 178 4 (POTENTIAL).
FT DOMAIN 179 208 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 209 229 5 (POTENTIAL).
FT DOMAIN 230 271 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 272 292 6 (POTENTIAL).
FT DOMAIN 293 305 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 307 327 7 (POTENTIAL).
FT DOMAIN 328 328 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 18 18 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 114 196 BY SIMILARITY.
SQ SEQUENCE 328 AA; 37648 MW; ADD7E25A41EACD0A CRC64;
Query Match 14.9%; Score 60.5; DB 1; Length 328;
Best Local Similarity 36.4%; Pred. No. 13;
Matches 20; Conservative 6; Mismatches 28; Indels 1; Gaps 1;
QY 13 EIIRLGSKNADGKLSVFGVLFDQDCANLFEAL-VCTLKAAARRKKIVTVAGELL 66
::: ||| | || | || | || | || | || | : |||||
Db 59 KLHRWTOKKGKKLSKMKVKLLKHLTLANLETLIYNPLDGWNITVQWYAGELL 113
RESULT 13
ARGB_BUCAI STANDARD; PRT; 257 AA.
AC PS7157;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acetylglutamate kinase (EC 2.7.2.8) (NAG kinase) (AGK) (N-acetyl-L-
DE glutamate 5-phosphotransferase).
GN ARGAB OR BU049.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;

```

```

DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine/threonine-protein kinase Pim-3 (EC 2.7.1.37) (Pim-1).
GN PIM3 OR PIM1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A., AND PHOSPHORYLATION SITES.
RX MEDLINE=97256766; PubMed=9099695;
RA Palaty C.K., Kaimar G., Tai G., Oh S., Amankawa L., Affolter M.,
RA Aebbersold R., Pelech S.L.;
RT "Identification of the autophosphorylation sites of the Xenopus
RT laevis Pim-1 proto-oncogene-encoded protein kinase.";
RL J. Biol. Chem. 272:10514-10521(1997).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- PIM: Autophosphorylated.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC PIM SUBFAMILY.
CC -!- CAUTION: Was originally (Ref.1) called Pim-1 but seems to
CC represent the pim-3 isoform.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L29495; AAA85389.1;
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
KW DOMAIN 40 291 PROTEIN KINASE.
FT NP_BIND 46 54 ATP (BY SIMILARITY).
FT BINDING 69 69 ATP (BY SIMILARITY).
FT ACT_SITE 168 168 BY SIMILARITY.
FT MOD_RES 4 4 PHOSPHORYLATION (AUTO-).
FT MOD_RES 190 190 PHOSPHORYLATION (AUTO-).
FT MOD_RES 205 205 PHOSPHORYLATION (AUTO-).
FT MOD_RES 205 205 PHOSPHORYLATION (AUTO-).
SQ SEQUENCE 323 AA; 36964 MW; AB4DD61E7A99A38F CRC64;

Query Match 15.0%; Score 61; DB 1; Length 323;
Best Local Similarity 32.6%; Pred. No. 11;
Matches 29; Conservative 13; Mismatches 27; Indels 20; Gaps 6;

QY 2 NVEH-EVNL-----VEEIHRLGSKNADGKLSVFGVLFQDDRCANLFEALVGTILK 51
DB 1 NMEHLPVKILQPVKVDKEFEKVKVQGVSWASG-----GEGTVISDRIAD--GQPVAVKH 70
QY 52 AAKRRKIVTYAGELLQGVDDVDVILQ 80
DB 71 VAKER--VTEWG--TLNGVMVPLEIVLTK 95

RESULT 11
OPPF_HAEIN
ID OPFF_HAEIN STANDARD; PRT; 332 AA.
AC P45051;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Oligopeptide transport ATP-binding protein oppf.
GN OPFF OR H1120.

OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC FOR OLIGOPEPTIDES. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE
CC TRANSPORT SYSTEM (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32792; AAC22774.1;
DR TIGR; H1120;
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transportr.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transportr; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW Peptide transport; Transport; Inner membrane; ATP-binding;
KW Complete proteome.
FT NP_BIND 56 63 ATP (POTENTIAL).
FT BINDING 332 AA; 36760 MW; 45B6A3085431E3D9 CRC64;
SQ SEQUENCE 332 AA; 36760 MW; 45B6A3085431E3D9 CRC64;

Query Match 15.0%; Score 61; DB 1; Length 332;
Best Local Similarity 33.3%; Pred. No. 11;
Matches 22; Conservative 4; Mismatches 26; Indels 14; Gaps 1;

QY 1 MNVEHEVNLVVEIHRGSKNADGKLSVFGVLFQDDRCANLFEALVGTILKAARRKIYV 60
DB 1 MIVSNKKELLEVNH-----LGVSFKNKDKSLFFAKPQTLKAVKDVSKL 46
QY 61 YAGELL 66
DB 47 YAGETL 52

RESULT 12
GRHR_HORSE
ID GRHR_HORSE STANDARD; PRT; 328 AA.
AC Q18821;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadotropin-releasing hormone receptor (GNRH-R).
GN GNRH.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;

```



```

OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: TO M.JANNASCHII MJ0977.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67526; AAB98830.1; -
DR TIGR; MJ0831; -
DR InterPro; IPR003156; DHHA1.
DR InterPro; IPR001667; Ppesterase.
DR Pfam; PF01368; DHH; 1.
DR Pfam; PF02272; DHHA1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 432 AA; 48835 MW; 0EB95D188D2B41F3 CRC64;

Query Match 16.0%; Score 65; DB 1; Length 432;
Best Local Similarity 26.9%; Pred. No. 5.3;
Matches 18; Conservative 15; Mismatches 20; Indels 14; Gaps 3;

QY 3 VEHEVN---LLVEIHRGSKNADGKLSVKGVLFDQDRCANL-----FEALVGTGK 51
DB 269 IEHKVRDAFLLEMLNVAVG---RNLFAVGIGICLEDDECEICIGNQLWEYKKNLNEK 325
QY 52 AAKRRKI 58
DB 326 SVKLKLL 332

RESULT 3
VG17_BPP2A STANDARD; PRT; 174 AA.
AC P08389;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Early protein GP17.
GN 17.
OS Bacteriophage P2A.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC phi-29-like viruses.
OX NCBI_TaxID=10757;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87031575; PubMed=3095189;
RA Paces V., Vicek C., Urbanek P., Hostomsky Z.;
RT "Nucleotide sequence of the right early region of Bacillus subtilis
RT phage P2A completes the 19366-bp sequence of P2A genome. Comparison
RT with the homologous sequence of phage phi 29.";
RL Gene 44:115-120(1986).
CC -1- FUNCTION: THIS PROTEIN IS INVOLVED IN THE REPLICATION OF THE

```

```

CC PHAGE DNA.
CC -1- SIMILARITY: BELONGS TO THE PODOVIRUSES GP17 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M11813; AAA88499.1; -
DR PIR; F29004; WRBP70.
KW Early protein; DNA replication.
SQ SEQUENCE 174 AA; 20022 MW; 08FBC8F91D4B4224 CRC64;

Query Match 15.9%; Score 64.5; DB 1; Length 174;
Best Local Similarity 32.4%; Pred. No. 2.2;
Matches 22; Conservative 12; Mismatches 31; Indels 3; Gaps 2;

QY 9 LLVEIHRGSKNADGKLSVKGVLFDQDRCANL- FEALVGTGKAAKRRKIYVYAGELL 67
DB 62 LLEVTVVEELSGSDNDLDEIDSELYDEAASDVPHETIVKVFADK--SIVTFNGERL 119
QY 68 QGVHDDVD 75
DB 120 HYVNVVDV 127

RESULT 4
HV21_HUMAN STANDARD; PRT; 146 AA.
AC P06331;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-II region ARH-77 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85205332; PubMed=3922855;
RA Kudo A., Ishihara T., Nishimura Y., Watanabe T.;
RT "A cloned human immunoglobulin heavy chain gene with a novel direct-
RT repeat sequence in 5' flanking region.";
RL Gene 33:181-189(1985).
DR PIR; A02101; G1HUH2.
DR HSSP; P01825; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR Immunoglobulin V region; Signal.
DR SIGNAL 1 19
FT CHAIN 20 146 IG HEAVY CHAIN V-II REGION ARH-77.
FT DOMAIN 20 117 V SEGMENT.
FT DOMAIN 118 127 D SEGMENT.
FT DOMAIN 128 146 J SEGMENT.
FT DISULFID 42 115 BY SIMILARITY.
FT NON_TER 146 146
SQ SEQUENCE 146 AA; 16228 MW; 8D7FD52BB218171F CRC64;

Query Match 15.8%; Score 64; DB 1; Length 146;
Best Local Similarity 35.4%; Pred. No. 2.1;
Matches 23; Conservative 8; Mismatches 30; Indels 4; Gaps 2;

QY 11 VEETHRLGSKNADGKLSVKGVLFDQDRCANL- FEALVGTGKAAKRRKIYVYAGELL 70
DB 68 IGEINHSGSTNY--KTSLSKSRVTISLDYTSKNLFSKLSSTVTAAD--TAVYICARGLRG 123
QY 71 HDDVD 75

```


A:Residues: 1-2535 <BEV>
A:Cross-references: EMBL:AL035440
A:Experimental source: cultivar Columbia; BAC clone F10M23
C:Genetics:
A:Map position: 4
A:Introns: 395/1; 613/3; 777/1; 826/3; 1945/2; 1993/3; 2061/3; 2129/3; 2178/3; 2295/3; 2
A:Note: F10M23.350

Query Match 15.5%; Score 63; DB 2; Length 2535;
Best Local Similarity 30.0%; Pred. No. 2.4e+02;
Matches 27; Conservative 13; Mismatches 32; Indels 18; Gaps 3;

QY 2 NVEHEVN-----LLVEEIHRLGSKNADGKLSVKFGVLFQDDRCANLFEALVGTLLKAAK 54
Db 477 NVMEVSSFFPDSQVLLIVLKLSDGSSGTOKLSLK-----REAELDSGLVGRKKRIK 528

QY 55 RRK---IVTYAGELLQGVHDDVDIVLLQD 81
Db 529 RSEKDVLEEEAVDIVIGGVGSKDIFLAED 558

Search completed: January 17, 2003, 02:17:36
Job time : 52 secs

T29652
inward rectifier potassium channel homolog - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T29652
R:Minx, P.; Le, T.T.
submitted to the EMBL Data Library, May 1996
A>Description: The sequence of C. elegans cosmid M02A10.
A:Accession: T29652
A:Reference number: Z20659
A>Status: preliminary; translated from GB/EMBL/DDDBJ
A:Molecule type: DNA
A:Residues: 1-514 <MIN>
A:CROSS-references: EMBL:U58730; PIDN:AAC48058.1; GSPDB:GN00028
A:Experimental source: strain Bristol N2; clone M02A10
C:Genetics:
A:Gene: irk-2
A:Map position: X
A:Introns: 87/3; 106/2; 151/3; 186/2; 210/3; 237/3; 268/3; 291/2; 334/1; 363/2; 404/3

Query Match 15.5%; Score 63; DB 2; Length 514;
Best Local Similarity 49.1%; Pred. No. 39;
Matches 26; Conservative 4; Mismatches 13; Indels 10; Gaps 4;

QY 19 SKNA-----DGKLSVKGVLFQDDRCANLFEALVGTLLKAARKRKIVTYAGELL 66
||||| ||||| : | : | : | : | : | : | : | : | : | : | : | : |
Db 234 SRNAVICLRDGLKCLLFRV--GDMRKSSLAEAHV-RLQMIRK-CVTYECELL 281

RESULT 14
S76690
hypothetical protein - Synecocystis sp. (strain PCC 6803)
C:Species: Synecocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S76690
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, K., 109-136, 1996
A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis s.

A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76690
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-811 <KAN>
A:CROSS-references: EMBL:D64004; GB:AB001339; NID:g1001701; PIDN:BAA10634.1; PID:g101
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

. Query Match 15.5%; Score 63; DB 2; Length 811;
. Best Local Similarity 22.1%; Pred. No. 66;
Matches 15; Conservative 18; Mismatches 31; Indels 4; Gaps 1;

QY 5 HEVNLLVEIHLRGSKNADGKLVSKEVLVFQDDRCANLFEALVGTLK----AAKKRKIVT 60
| | | | : : : : | : : | : | : | : | : | : | : | : | : | : | : |
Db 291 HLNRITLPFOAHQVAENSNNLLDKYNLTVDKEATKNLAEKAVLAMENKVSAEKNVIV 350

QY 61 YAGELLQ 68
|||::| |

Db 351 RAGEVITQ 358

RESULT 15
T04824
hypothetical protein FLOW23.350 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C:Accession: T04824
R:Bevan, M.; Lecharny, A.; Chedford, F.; Krivitzky, M.; Kreis, M.; Hoheisel, J.; Mewes
submitted to the Protein Sequence database, February 1999
A:Reference number: Z15385
A:Accession: T04824
A:Molecule type: DNA

[illegible]

Query Match 23.8%; Score 96.5; DB 2; Length 162;
Best Local Similarity 32.9%; Pred. No. 0.002;
Matches 26; Conservative 16; Mismatches 34; Indels 3; Gaps 2;

QY 2 NVEHEVNLLVEEIHRLGSKNADG-KLSVKFGVLQDDRCANLFEALVGTTLKAARRKIVT 60
DB 66 HVHREMLTLCWEVDYKQKEGSPRITFGRLF--TIYVNISDKVVGVTLLRARRKKMD 123
QY 61 YAGELLQGVHDDVDIVLL 79
DB 124 FEGEMLFQKRDDHVIILL 142

RESULT 3
T24421
hypothetical protein T04A8.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24421
R:Palmer, S.
submitted to the EMBL Data Library, August 1994
A:Reference number: Z19889
A:Accession: T24421
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-289 <MIL>
A:Cross-references: EMBL:Z35663; PIDN:CAA84722.1; GSPDB:GN00021; CESP:T04A8.4
A:Experimental source: clone T04A8
C:Genetics:
A:Gene: CESP:T04A8.4
A:Map position: 3
A:Introns: 26/2; 58/3; 102/3; 191/2; 223/3

Query Match 20.1%; Score 81.5; DB 2; Length 289;
Best Local Similarity 33.7%; Pred. No. 0.18;
Matches 29; Conservative 14; Mismatches 26; Indels 17; Gaps 4;

QY 1 MNVEHEVNLLVEEIHRLGSKNADGK---LSVKFGVLQDDRCANLFEALVGTTLKAA 53
DB 61 VHCREILFLCETI-----DSNADGSEEPKVKFGKLF-----NIYFYSKLVGLMIRA 110
QY 54 KRRKIVTYAGELLQGVHDDVDIVLL 79
DB 111 RYGLVHFEGEMLYQRODDEKIITML 136

RESULT 4
E90541
hypothetical protein MYPV_2370 [Imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: E90541
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Sanson, D.; Gallisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: E90541
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-318 <KOR>
A:Cross-references: GB:AL445566; PID:gl4089650; PIDN:CAC13410.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPV_2370
A:Genetic code: SGC3
C:Superfamily: phosphate acetyltransferase

Query Match 17.0%; Score 69; DB 2; Length 318;
Best Local Similarity 36.2%; Pred. No. 4.9;
Matches 29; Conservative 6; Mismatches 25; Indels 20; Gaps 4;

QY 6 EVNLLVEEIHRLGSKNADGKLSVKFGVLQDDRCANLFEALVGTTLKAAK 54

DB 44 EYSLIVE-----SKSDYKGFNFVLLDQDQKQYSEFCODLFESRKGRDLSLVOKALK 96
QY 55 RRRKIVTYAGELLQGVHDDV 74
DB 97 TRPF--YAMMLLKGFFDGV 114

RESULT 5
F81343
probable integral membrane zinc-metalloprotease Cj0723c [Imported] - Campylobacter jejuni
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: F81343
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chli
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: F81343
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-395 <PAR>
A:Cross-references: GB:AL139076; GB:AL111168; NID:g6968128; PIDN:CAB72997.1; PID:g696
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj0723c

Query Match 16.9%; Score 68.5; DB 2; Length 395;
Best Local Similarity 26.2%; Pred. No. 7.1;
Matches 21; Conservative 19; Mismatches 29; Indels 11; Gaps 3;

QY 2 NVEHEVNLLVEE-----IHLGSKNADGKLSVKFGVLQDDRCANLFEALVGTTLKAAK 54
DB 200 NLLKISSLLMKQCGFSANGVYVIDASKRDLNAYEGGLFKSRVV-LFDTL---LKALN 255
QY 55 RRRKIVTYAGELLQGVHDDV 74
DB 256 ERELLANVLGHELGHFVHKDI 275

RESULT 6
T04871
hypothetical protein F28A21.200 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Sep-1999
C:Accession: T04871
R:Bevan, M.; Mueller, M.W.; Muendlein, A.; Felber, R.; Bancroft, I.; Mewes, H.W.; May
submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15387
A:Accession: T04871
A:Molecule type: DNA
A:Residues: 1-530 <BEV>
A:Cross-references: EMBL:AL035526
A:Experimental source: cultivar Columbia; BAC clone F28A21
C:Genetics:
A:Map position: 4
A:Introns: 193/2; 239/1; 456/3
A:Note: F28A21.200
C:Superfamily: natural resistance-associated macrophage protein 1

Query Match 16.3%; Score 66; DB 2; Length 530;
Best Local Similarity 29.5%; Pred. No. 19;
Matches 23; Conservative 11; Mismatches 36; Indels 8; Gaps 3;

QY 3 VEHEVNLLVEEIHRLGSKNADGKLSVKFGVLQDDRCANLFEALVGTTLKAAK 55
DB 303 VSEMINLFTVAFKGFYGTQKQADSIGLVNAGVYLQEKYGGVFPILYIWIGIGLLAAGS 362
QY 56 RKIV-TYAGELLQGVHDDV 72
DB 363 STITGTAGQFTMEGFLD 380

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 02:02:38 ; Search time 49 Seconds

(without alignments)
158.916 Million cell updates/sec

Title: US-09-648-310-2

Perfect score: 406

Sequence: 1 MNVEHVNLLVEEIRHRLGSK.....AGELLQGVHDDVDIVLLQD 81

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	286	70.4	161	2 T04981	hypothetical prote
2	96.5	23.8	162	2 T21860	hypothetical prote
3	81.5	20.1	289	2 T24421	hypothetical prote
4	69	17.0	318	2 E90541	hypothetical prote
5	68.5	16.9	395	2 F81343	probable integral
6	66	16.3	530	2 T04871	hypothetical prote
7	66	16.3	1404	2 E36788	hypothetical prote
8	65	16.0	432	2 G64403	hypothetical prote
9	64.5	15.9	174	1 WRBP70	early protein gp17
10	64	15.8	146	1 GLHDH2	Ig heavy chain pre
11	64	15.8	599	2 H72336	conserved hypothet
12	63.5	15.6	496	2 F84784	probable glucosyl
13	63	15.5	514	2 T29652	inward rectifier p
14	63	15.5	811	2 S76690	hypothetical prote
15	63	15.5	2535	2 T04824	hypothetical prote
16	62.5	15.4	491	2 C84784	probable glucosyl
17	62.5	15.4	496	2 D84784	probable glucosyl
18	62.5	15.4	565	2 G64532	methyl-accepting c
19	62.5	15.4	902	2 S33918	dynam-in-like prote
20	62	15.3	268	2 AE2515	hypothetical prote
21	62	15.3	1242	2 S14201	probable adenylate
22	61.5	15.1	637	1 WQSC2M	phosphotransferase
23	61.5	15.1	988	2 AE2697	glutamate-ammonia-
24	61.5	15.1	1010	2 E97479	hypothetical prote
25	61	15.0	217	2 AH0561	potential acraB op
26	61	15.0	332	2 B64184	oligopeptide trans
27	61	15.0	556	2 A81931	probable adhesin N
28	61	15.0	635	2 T15904	hypothetical prote
29	60.5	14.9	564	2 F71975	methyl-accepting c

30 60.5 14.9 573 2 F96560
31 60.5 14.9 1870 2 D88486
32 60 14.8 257 2 D84935
33 60 14.8 659 2 D84286
34 60 14.8 731 2 D82249
35 60 14.8 824 2 T44037
36 59.5 14.7 147 1 MCBY
37 59.5 14.7 310 2 S41055
38 59.5 14.7 351 2 H72328
39 59.5 14.7 682 2 T50840
40 59.5 14.7 683 2 T50839
41 59.5 14.7 775 2 E69435
42 59 14.5 287 2 G64504
43 59 14.5 506 2 AH0183
44 59 14.5 637 2 C91188
45 59 14.5 637 2 C86035

ALIGNMENTS

RESULT 1

T04981

hypothetical protein T161L130 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 24-Nov-1999

C;Accession: T04981

R;Bayan, M.; Obermaier, B.; Deutschenbaur, S.; Piravandi, E.; Hoheisel, J.; Mewes, H.

Submitted to the Protein Sequence Database, November 1998

A;Reference number: Z15393

A;Accession: T04981

A;Molecule type: DNA

A;Residues: 1-161 <BEV>

A;Cross-references: EMBL:AL031394

A;Experimental source: cultivar Columbia; BAC clone T161L

C;Genetics:

A;Map position: 4

A;Introns: 93/3

A;Note: T161L130

C;Superfamily: Arabidopsis thaliana hypothetical protein T161L130

Query Match 70.4%; Score 286; DB 2; Length 161;
Best Local Similarity 70.5%; Pred. No. 1.5e-24;
Matches 55; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 1 MNVEHVNLLVEEIRHRLGSKNADGKLSVKFGLFODRCANLFEALVGTLLKAARKKIWT 60

Db 67 MNVEEIQKLEEEIRHRLGSRDTGSKYKVTFGVLFNDRCANIFEALVGTLLRAAKRKIVA 126

QY 61 YAGELLQGVHDDVDIVL 78

Db 127 FEGLLQGVHDKVEITL 144

RESULT 2

T21860

hypothetical protein F36F2.1 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T21860

R;Cottage, A.

Submitted to the EMBL Data Library, November 1996

A;Reference number: Z19479

A;Accession: T21860

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-162 <WIL>

A;Cross-references: EMBL:Z81532; PIDN: CAB04325.1; GSPDB: GNO00019; CESP: F36F2.1

A;Experimental source: clone F36F2

C;Genetics:

A;Gene: CESP:F36F2.1

A;Map position: 1

A;Introns: 9/1; 28/1; 64/1; 108/3

Fri Jan 17 09:17:31 2003

us-09-648-310-2.ra1

Page 6

Search completed: January 17, 2003, 02:18:24
Job time : 37 secs

FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4164
LENGTH: 128
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4164

Query Match 13.4%; Score 54.5; DB 4; Length 128;
Best Local Similarity 30.6%; Pred. No. 8.7;
Matches 26; Conservative 13; Mismatches 35; Indels 11; Gaps 3;

QY 4 EHEVNLVEEIHRLGSK-----KNADGKLSVKFGVLFQDDRCANLFEALVGTLLKAARR 56
Db 22 QVELEKLLKENEELKAEKALSQMKNETRSMNLNESGLENFDDQIVNI--LVNT-DAEKTR 77
QY 57 KIVTYAGELLQGVHDDVDVLLQD 81
Db 78 KNVESPTNLLNQMKSVNKEALRQD 102

RESULT 13
US-09-058-489-18
Sequence 18, Application US/09058489
Patent No. 6103886
GENERAL INFORMATION:
APPLICANT: Whitehead Institute for Biomedical Research
APPLICANT: Lahn, Bruce
APPLICANT: Page, David
TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
FILE REFERENCE: WHI97-08pA
CURRENT APPLICATION NUMBER: US/09/058,489
CURRENT FILING DATE: 1998-04-10
EARLIER APPLICATION NUMBER: 60/041,877
EARLIER FILING DATE: 1997-04-11
NUMBER OF SEQ ID NOS: 91
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 18
LENGTH: 660
TYPE: PRT
ORGANISM: Human
US-09-058-489-18

Query Match 13.4%; Score 54.5; DB 3; Length 660;
Best Local Similarity 27.1%; Pred. No. 69;
Matches 16; Conservative 12; Mismatches 24; Indels 7; Gaps 2;

QY 2 NVEHEVNL-----LVEEIHRLGSKNADGKLSVKFGVLFQDDRCANLFEALVGTLLKAARR 55
Db 507 NVHRVINFDPSPDIEYVHRIGTRGVNGL--ATSFNEKNMNITKDLLDLLEAKRQ 563

RESULT 14
US-09-058-489-91
Sequence 91, Application US/09058489
Patent No. 6103886
GENERAL INFORMATION:
APPLICANT: Whitehead Institute for Biomedical Research
APPLICANT: Lahn, Bruce
APPLICANT: Page, David
TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
FILE REFERENCE: WHI97-08pA
CURRENT APPLICATION NUMBER: US/09/058,489
CURRENT FILING DATE: 1998-04-10
EARLIER APPLICATION NUMBER: 60/041,877

EARLIER FILING DATE: 1997-04-11
NUMBER OF SEQ ID NOS: 91
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 91
LENGTH: 660
TYPE: PRT
ORGANISM: Human
US-09-058-489-91

Query Match 13.4%; Score 54.5; DB 3; Length 660;
Best Local Similarity 27.1%; Pred. No. 69;
Matches 16; Conservative 12; Mismatches 24; Indels 7; Gaps 2;

QY 2 NVEHEVNL-----LVEEIHRLGSKNADGKLSVKFGVLFQDDRCANLFEALVGTLLKAARR 55
Db 507 NVHRVINFDPSPDIEYVHRIGTRGVNGL--ATSFNEKNMNITKDLLDLLEAKRQ 563

RESULT 15
US-08-952-127-12
Sequence 12, Application US/08952127
Patent No. 6211336
GENERAL INFORMATION:
APPLICANT: Shiloh, Yosef
APPLICANT: Tagle, Danilo A.
APPLICANT: Collins, Francis S.
TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates
STREET: 30500 No. 6211336thwestern Hwy., Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: U.S.
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/952,127
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,995
REFERENCE/DOCKET NUMBER: 2290.00029
TELECOMMUNICATION INFORMATION:
TELEPHONE: 810-539-5050
TELEFAX: 810-539-5055
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3066 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Mus musculus
US-08-952-127-12

Query Match 13.4%; Score 54.5; DB 4; Length 3066;
Best Local Similarity 26.6%; Pred. No. 4.7e+02;
Matches 25; Conservative 16; Mismatches 28; Indels 25; Gaps 5;

QY 8 NLLVEEIHRLGSK-----NADGKLSVKFGVLFQDDRCANLFEALVGTLLKAA----- 53
Db 317 DLLVNEISHIGSRKYSYSSGRNIAVKNLIDLMDADICQLDFADTRSVESIQSVVTQRES 376
QY 54 -----KRRKIVTYAGELL---LQGVHDDVDIV 77
Db 377 TDYSVPCKRRKI-DVGWEVINKDYLOKQSQSDFLV 409

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,503
; FILING DATE: 5-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-503-2

Query Match 13.8%; Score 56; DB 2; Length 2206;
Best Local Similarity 36.2%; Pred. No. 2e+02;
Matches 17; Conservative 6; Mismatches 8; Indels 16; Gaps 2;

QY 47 VGTLLKAAK-----RRKIVTYA---GELLQGVHDDVDIV 77
| : : : : : | : : : : : | : : : : : | : : : : : |
Db 1553 VPTIRAAKVGPGFDYAVAMAKRNIVTATTSKGEFTMLGVHONVAIL 1599

```

```

RESULT 10
US-08-465-250-2
; Sequence 2, Application US/08465250
; Patent No. 6136570
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release 1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,250
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-E-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

US-08-465-250-2

```

Query Match 13.8%; Score 56; DB 4; Length 2206;
Best Local Similarity 36.2%; Pred. No. 2e+02;
Matches 17; Conservative 6; Mismatches 8; Indels 16; Gaps 2;

QY 47 VGTLLKAAK-----RRKIVTYA---GELLQGVHDDVDIV 77
| : : : : : | : : : : : | : : : : : | : : : : : |
Db 1553 VPTIRAAKVGPGFDYAVAMAKRNIVTATTSKGEFTMLGVHONVAIL 1599

```

```

RESULT 11
US-08-203-905B-14
; Sequence 14, Application US/08203905B
; Patent No. 5646249
; GENERAL INFORMATION:
; APPLICANT: KAYE, FEDERIC J.
; APPLICANT: OTTERSON, GREGORY A.
; TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF A
; TITLE OF INVENTION: NOVEL CHAPERONE PROTEIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON & BEAR
; STREET: 620 NEWPORT CENTER DRIVE, SIXTEENTH FLOOR
; CITY: NEWPORT BEACH
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,905B
; FILING DATE: February 28, 1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: KIRKPATRICK, ANITA M.
; REGISTRATION NUMBER: 32,617
; REFERENCE/DOCKET NUMBER: NIH089.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-203-905B-14

```

```

Query Match 13.7%; Score 55.5; DB 1; Length 472;
Best Local Similarity 27.0%; Pred. No. 33;
Matches 27; Conservative 12; Mismatches 30; Indels 31; Gaps 3;

QY 7 VNLIVVEIHLGSKNADGKL-----SVKFGVLFQDDRCAN 41
| : : : : : | : : : : : | : : : : : | : : : : : |
Db 310 VLLTVEENDSQKPNADSKLPEDQLTPGQHVNVRFRGLSDSTSAKQVLFETVSRK 369

QY 42 LFEALVGLTKAAKRNIVTYAGELLQGVHD--DVDIVLL 79
| : : : : : | : : : : : | : : : : : | : : : : : |
Db 370 LF-----NTLNEDLFQKILVPIQOVLKGLDKTEIDVYL 405

```

```

RESULT 12
US-09-134-001C-4164
; Sequence 4164, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

```

; LENGTH: 2052 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
US-09-045-201A-2

Query Match 13.8%; Score 56; DB 3; Length 2052;

Best Local Similarity 26.0%; Pred. No. 1.8e+02;

Matches 19; Conservative 15; Mismatches 23; Indels 16; Gaps 4;

QY 14 IHRIGSKNADGKLSVKFGV-----LFQDDRCANLFEALVGLTKAAKRKIYTYAGELLQ 68

Db 1870 VIRIGYKNSQNTTEKKDLLVWENLFYGRKMAQVFD-LKGSLL---RNRNVTDTGK---- 1921

QY 69 GVHDDVDIVLQD 81

Db 1922 ---ESCDVLLDE 1931

RESULT 7

US-09-619-062-2

; Sequence 2, Application US/09619062

; Patent No. 6406875

; GENERAL INFORMATION:

; APPLICANT: Shisheva, Assia

; TITLE OF INVENTION: No. 6406875el Mammalian Putative

; TITLE OF INVENTION: Phosphatidylinositol-4-Phosphate-5-Kinase

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Harness, Dickey & pierce, P.L.C.

; STREET: P.O. Box 828

; CITY: Bloomfield Hills

; STATE: Michigan

; COUNTRY: U.S.A.

; ZIP: 48303

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/619,062

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/045,201

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Smith, DeAnn F

; REFERENCE/DOCKET NUMBER: 4981-098431

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (248) 641-1600

; TELEFAX: (248) 641-0270

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2052 amino acids

; TYPE: amino acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

US-09-619-062-2

Query Match

Best Local Similarity 13.8%; Score 56; DB 4; Length 2052;

Matches 19; Conservative 15; Mismatches 23; Indels 16; Gaps 4;

QY 14 IHRIGSKNADGKLSVKFGV-----LFQDDRCANLFEALVGLTKAAKRKIYTYAGELLQ 68

Db 1870 VIRIGYKNSQNTTEKKDLLVWENLFYGRKMAQVFD-LKGSLL---RNRNVTDTGK---- 1921

QY 69 GVHDDVDIVLQD 81

Db 1922 ---ESCDVLLDE 1931

RESULT 8

US-07-852-260-2

; Sequence 2, Application US/07852260

; Patent No. 5525715

; GENERAL INFORMATION:

; APPLICANT: Racanello, Vincent

; APPLICANT: Tatem, Joanne M.

; APPLICANT: Weeks-Levy, Carolyn L.

; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM

; TITLE OF INVENTION: CDNA

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham

; STREET: 30 Rockefeller Plaza

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10112

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/852,260

; FILING DATE: 19920619

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 36607-B-PCT-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 977-9550

; TELEFAX: (212) 664-0525

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2206 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-07-852-260-2

Query Match

Best Local Similarity 13.8%; Score 56; DB 1; Length 2206;

Matches 17; Conservative 6; Mismatches 8; Indels 16; Gaps 2;

QY 47 VGTLLKAAK-----RRKIVTYA---GELLQGVHDDVDIV 77

Db 1553 VPTIRAAKVGPGFDYAVAMAKRNIYVATTTSKGEFTMLGVHDNVAIL 1599

RESULT 9

US-08-461-503-2

; Sequence 2, Application US/08461503

; Patent No. 5834302

; GENERAL INFORMATION:

; APPLICANT: Racanello, Vincent

; APPLICANT: Tatem, Joanne M.

; APPLICANT: Weeks-Levy, Carolyn L.

; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES

; TITLE OF INVENTION: FROM CDNA

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10112

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

```
Query Match      15.0%; Score 61; DB 4; Length 323;
Best Local Similarity 32.6%; Pred. No. 3.8;
Matches 29; Conservative 13; Mismatches 27; Indels 20; Gaps 5;

QY 2 NVEHEVNLL-----LVEEHLRLGSKNADGKLSVKFGVLFQDDRCANLFEALVGTLLK 51
Db 17 NMEHLPVKILQPVKDKPEFKYQVGSVVASG----GFGTVYSDSRIAD--GQPVAVKH 70

QY 52 AAKRRKIYTAGELLQGVHDDVDIVLLQ 80
Db 71 VAKER--VTEWG--TLNGVMPLEIVLLK 95

RESULT 3
US-09-058-489-15
; Sequence 15, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WH197-08pA
; CURRENT APPLICATION NUMBER: US/09/058,489
; EARLIER FILING DATE: 1998-04-10
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Human
US-09-058-489-15

Query Match      14.4%; Score 58.5; DB 3; Length 662;
Best Local Similarity 28.8%; Pred. No. 20;
Matches 17; Conservative 12; Mismatches 23; Indels 7; Gaps 2;

QY 2 NVEHEVNLL-----LVEEHLRLGSKNADGKLSVKFGVLFQDDRCANLFEALVGTLLK 55
Db 509 NVKHVINFLPSDIEEYVHRIGTRGVNGL--ATSPFNERNINITKDLLLVEAK 565

RESULT 4
US-09-058-489-16
; Sequence 16, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WH197-08pA
; CURRENT APPLICATION NUMBER: US/09/058,489
; EARLIER FILING DATE: 1998-04-10
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Human
US-09-058-489-16

Query Match      14.4%; Score 58.5; DB 3; Length 662;
Best Local Similarity 28.8%; Pred. No. 20;
Matches 17; Conservative 12; Mismatches 23; Indels 7; Gaps 2;
```

```
QY 2 NVEHEVNLL-----LVEEHLRLGSKNADGKLSVKFGVLFQDDRCANLFEALVGTLLK 55
Db 509 NVKHVINFLPSDIEEYVHRIGTRGVNGL--ATSPFNERNINITKDLLLVEAK 565

RESULT 5
US-09-134-001C-2873
; Sequence 2873, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCI
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2873
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2873

Query Match      13.8%; Score 56; DB 4; Length 208;
Best Local Similarity 28.3%; Pred. No. 10;
Matches 22; Conservative 14; Mismatches 22; Indels 20; Gaps 4;

QY 10 LVEEHLRLGSKNADGKLSVKFGVLFQDDRCANLFEALVGTLLKAAKRRKIVTVAG 63
Db 29 MVRCHSI--REARRLNVASSAVNRQILKLEDEMGATFLDRLPGLR-----VTAAG 79

QY 64 ELLQGVHDDVDIVLLQD 81
Db 80 EILTRHI-----TILLQD 92

RESULT 6
US-09-045-201A-2
; Sequence 2, Application US/09045201A
; Patent No. 6110718
; GENERAL INFORMATION:
; APPLICANT: Shisheva, Assia
; TITLE OF INVENTION: No. 6110718el Mammalian Putative
; TITLE OF INVENTION: Phosphatidylinositol-4-Phosphate-5-Kinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,201A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Dean F
; REFERENCE/DOCKET NUMBER: 4981-098431
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 641-1600
; TELEFAX: (248) 641-0270
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 02:04:14 ; Search time 36 Seconds
(without alignments)
66.202 Million cell updates/sec

Title: US-09-648-310-2
Perfect score: 406
Sequence: 1 MNVEHEVNLVEEHRGSK.....AGELLQGVHDDVDIVLLQD 81

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/BACKFILES.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	61	15.0	323	4	US-09-237-543-6
2	61	15.0	323	4	US-09-644-450-6
3	58.5	14.4	662	3	US-09-058-489-15
4	58.5	14.4	662	3	US-09-058-489-16
5	56	13.8	208	4	US-09-134-001C-2873
6	56	13.8	2052	3	US-09-045-201A-2
7	56	13.8	2052	4	US-09-615-062-2
8	56	13.8	2206	1	US-07-852-260-2
9	56	13.8	2206	2	US-08-461-503-2
10	56	13.8	2206	4	US-08-465-250-2
11	55.5	13.7	472	1	US-08-203-905B-14
12	54.5	13.4	128	4	US-09-134-001C-4164
13	54.5	13.4	660	3	US-09-058-489-18
14	54.5	13.4	660	3	US-09-058-489-91
15	54.5	13.4	3066	4	US-08-952-127-12
16	54	13.3	129	4	US-08-981-527A-11
17	54	13.3	864	2	US-08-209-521-29
18	53.5	13.2	402	4	US-09-347-801-16
19	53.5	13.2	466	4	US-08-914-375C-60
20	53.5	13.2	966	1	US-08-571-758-2
21	53.5	13.2	966	1	US-08-909-984A-2
22	53.5	13.2	966	1	US-08-909-983-2
23	53	13.1	289	4	US-09-071-035-72
24	53	13.1	312	4	US-09-071-035-70
25	53	13.1	367	2	US-08-530-230-24
26	52.5	12.9	138	4	US-09-134-001C-2982
27	52.5	12.9	689	1	US-07-766-351-5

28	52.5	12.9	689	1	US-08-059-032-5	Sequence 5, Appli
29	52.5	12.9	689	5	PCT-US91-07290-5	Sequence 5, Appli
30	52.5	12.9	739	2	US-08-836-943-2	Sequence 2, Appli
31	52	12.8	489	1	US-08-095-726-8	Sequence 8, Appli
32	52	12.8	489	1	US-08-095-726-10	Sequence 10, Appli
33	52	12.8	489	1	US-08-096-043-8	Sequence 8, Appli
34	52	12.8	489	1	US-08-096-043-10	Sequence 10, Appli
35	52	12.8	489	1	US-08-096-623A-8	Sequence 8, Appli
36	52	12.8	489	1	US-08-096-623A-10	Sequence 10, Appli
37	52	12.8	510	4	US-09-291-922-22	Sequence 22, Appli
38	51.5	12.7	155	2	US-08-844-057-4	Sequence 4, Appli
39	51.5	12.7	155	4	US-09-006-730-4	Sequence 4, Appli
40	51.5	12.7	315	4	US-09-134-001C-4771	Sequence 4771, Ap
41	51.5	12.7	425	4	US-09-134-001C-5619	Sequence 5619, Ap
42	51.5	12.7	872	2	US-08-844-057-2	Sequence 2, Appli
43	51.5	12.7	872	4	US-09-006-730-2	Sequence 2, Appli
44	51	12.6	946	3	US-08-560-005-4	Sequence 4, Appli
45	51	12.6	946	4	US-09-418-540-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-237-543-6
; Sequence 6, Application US/09237543A
; Patent No. 6143540
; GENERAL INFORMATION:
; APPLICANT: Kapeller, Rosana
; TITLE OF INVENTION: NOVEL MOLECULES OF THE HKID-1-RELATED PROTEIN FAMILY
; FILE REFERENCE: 035800/175631
; CURRENT APPLICATION NUMBER: US/09/237,543A
; CURRENT FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 6
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-237-543-6

Query Match 15.0%; Score 61; DB 4; Length 323;
Best Local Similarity 32.6%; Pred. No. 3.8;
Matches 29; Conservative 13; Mismatches 27; Indels 20; Gaps 6;

QY 2 NVEH-EVNLL-----VEEHRGSKNADGKLSVKFGLFODDRCANLFEALVGLK 51
DB 17 NMEHLPVKILQPKVKDEPFKVKYGVGSVVASG-----GFGTVYSDSRIAD--GQPVAVKH 70
QY 52 AAKRRKIVYAGELLLQGVHDDVDIVLLQ 80
DB 71 VAKER--VTENG--TLNGVMVPLEIVLLK 95

RESULT 2
US-09-644-450-6
; Sequence 6, Application US/09644450
; Patent No. 6383791
; GENERAL INFORMATION:
; APPLICANT: Kapeller, Rosana
; TITLE OF INVENTION: NOVEL MOLECULES OF THE HKID-1-RELATED PROTEIN FAMILY
; FILE REFERENCE: 035800/175631
; CURRENT APPLICATION NUMBER: US/09/644,450
; CURRENT FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 6
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-644-450-6

```
; Patent No. US20010055759A1
; GENERAL INFORMATION:
; APPLICANT: Kathariou, Sophia
; APPLICANT: Lei, Xiang-Hei
; TITLE OF INVENTION: Serotype-Specific Probes for Listeria Monocytogenes
; FILE REFERENCE: A65378/RFT/DSS
; CURRENT APPLICATION NUMBER: US/09/115,150
; CURRENT FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Bacterial
US-09-115-150-4

Query Match      13.4%; Score 54.5; DB 10; Length 549;
Best Local Similarity 29.8%; Pred. No. 46;
Matches 25; Conservative 13; Mismatches 33; Indels 13; Gaps 4;

QY 4 EHEVNLLVEIHRGSKNADGKLSVFGVLFQDDRCANLFEALVGTLKAARKRKIV--TY 61
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 405 EEEINKLGYEVQNEYKNATGKDAVSNEYKR--RIDDSNPLISISKAPSLKKTDSIY 462
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

QY 62 A-----GELLQGVHDDVDIVLL 79
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 463 AKTTNEMQNYLLRGA---IPLVLL 483
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

RESULT 13
US-09-925-301-1158
; Sequence 1158, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1158
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (129)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (150)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1158

Query Match      13.4%; Score 54.5; DB 10; Length 639;
Best Local Similarity 25.9%; Pred. No. 56;
Matches 14; Conservative 18; Mismatches 17; Indels 5; Gaps 2;

QY 5 HEVNLLVEIHRGSKNADGKLSV-KFGVLFQDDRCANLFEALVGTLKAARKRK 57
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 51 YKVEIIQKLMIDGRNKGDKISDFEYVIFQEVKSSDIK---TFRKAINRK 100
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

RESULT 14
US-09-817-514A-2
; Sequence 2, Application US/09817514A
; Patent No. US20020078478A1
; GENERAL INFORMATION:
; APPLICANT: french-Constant, Richard
```

```
; APPLICANT: Bowen, David
; APPLICANT: Rocheleau, Thomas
; APPLICANT: Waterfield, Nicholas
; TITLE OF INVENTION: DNA SEQUENCES FROM PHOTORHABDUS LUMINESCENS
; FILE REFERENCE: 61645
; CURRENT APPLICATION NUMBER: US/09/817,514A
; CURRENT FILING DATE: 2000-03-26
; PRIOR APPLICATION NUMBER: US 60/191806
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 2516
; TYPE: PRT
; ORGANISM: Photorhabdus luminescens
US-09-817-514A-2

Query Match      13.4%; Score 54.5; DB 10; Length 2516;
Best Local Similarity 23.9%; Pred. No. 3.2e+02;
Matches 26; Conservative 15; Mismatches 33; Indels 35; Gaps 4;

QY 1 MNVEHEVNLL-----LVEETHRL-----GSKN-----ADGKLSYK 29
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 2329 LEVERTVSLAEVYAGLPKDNNGFFSLAQEIDKLVSQSGSAGSNNNLAFGAGTDIKTSLQ 2388
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

QY 30 FGVLFQDDRCANLFEALVGTLKAARKRKIVTYAGELLQGVHDDVDIVL 78
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 2389 ASVSPADLKIREDPASLGKIRRIKQISVTLP-----LLGPTQDVQAIL 2433
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

RESULT 15
US-09-788-657-16
; Sequence 16, Application US/09788657
; Patent No. US20020123149A1
; GENERAL INFORMATION:
; APPLICANT: Nicolaides, Nicholas
; APPLICANT: Sass, Philip
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Grasso, Luigi
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: Methods for generating hypermutable
; FILE REFERENCE: 01107.00097
; CURRENT APPLICATION NUMBER: US/09/788,657
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,336
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 859
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-788-657-16

Query Match      13.3%; Score 54; DB 10; Length 859;
Best Local Similarity 29.2%; Pred. No. 94;
Matches 14; Conservative 12; Mismatches 14; Indels 8; Gaps 2;

QY 26 LSVKFGVLFQDDRCANLFEALVGT-LKAARKRKIVTYAGEL 65
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 787 LSDSPGVMCPSRVQMFASRACRSKVMIGTALNASEMKKLTHMGEM 834
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

Search completed: January 17, 2003, 02:19:49
Job time : 73 secs
```

```

US-09-767-041-10
; Sequence 10, Application US/09767041
; Patent No. US20020055168A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Hilda
; TITLE OF INVENTION: STREPTOCOCCUS SUIS VACCINES AND DIAGNOSTIC TESTS
; FILE REFERENCE: 2183-4726
; CURRENT APPLICATION NUMBER: US/09/767,041
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: PCT/NL99/00460
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: EP98202465.5
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: EP98202467.1
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Streptococcus suis
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: ORF22
US-09-767-041-10

Query Match 13.4%; Score 54.5; DB 10; Length 239;
Best Local Similarity 24.8%; Pred. No. 16;
Matches 28; Conservative 15; Mismatches 35; Indels 35; Gaps

QY 1 MNVEHEVNLVVEIHRGLSKNAD-----GKL---SVKFGVLFODDRCANLF 43
Db 76 LSASGEMDLLVHQINRLISAGLDPPQVVEALTHYREHSKLLFLVLAKVDNLVKNRGLSKLV 135
QY 44 EALVGTL-----KAAKRKIVTYAGELLQGVHDDVDIVL 78
Db 136 GTVGLLNIRMVGEASAEGLLELQKARGHKHSVTAFAEEMKAGYDGRIVM 188

RESULT 11
US-09-825-882-12
; Sequence 12, Application US/09825882
; Patent No. US20020094551A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; TITLE OF INVENTION: T2R TASTE RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 078003/0279152/RXT
; CURRENT APPLICATION NUMBER: US/09/825,882
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,532
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/247,014
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-825-882-12

Query Match 13.4%; Score 54.5; DB 10; Length 309;
Best Local Similarity 42.9%; Pred. No. 22;
Matches 15; Conservative 8; Mismatches 11; Indels 1; Gaps

QY 32 VLFQDDRCANLFEALVGTLKAAKRKIVTYAGELL 66
Db 15 VLFVIGNFANGFIALVNSTIERVKRKQKI-SFADQIL 48

RESULT 12
US-09-115-150-4
; Sequence 4, Application US/09115150

```



```

; APPLICANT: Laura A. Rudolph-Owen
; APPLICANT: Kyle MacBeth
; TITLE OF INVENTION: NOVEL MOLECULES OF THE HKID-1-RELATED PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 35800/238856
; CURRENT APPLICATION NUMBER: US/09/971,791
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 09/644,450
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/237,543
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-971-791-6

Query Match      15.08; Score 61; DB 10; Length 323;
Best Local Similarity 32.6%; Pred. No. 3.5;
Matches 29; Conservative 13; Mismatches 27; Indels 20; Gaps 6;

Qy  2 NVEHL-EVNL-----VEIHRIGSKNADGKLSVFGVLFQDDRCANLFEALVGTILK 51
Db  17 NMEHL-PVKILQPVKVDKPEFKVQGVSVASG-----CFGIVSDRIAD--GQPVAVKH 70

Qy  52 AAKRRKIVTYAGELLQGVHDDVDIVLQ 80
Db  71 VAKER--VTENG--TLNGVMVPLEIVLK 95

RESULT 3
US-09-815-242-5045
; Sequence 5045, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5045
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae

Query Match      14.08; Score 57; DB 10; Length 506;
Best Local Similarity 32.5%; Pred. No. 20;
Matches 26; Conservative 10; Mismatches 26; Indels 18; Gaps 5;

Qy  7 VNLV--EEIHRIGSKNADGK---LSVKFGVLFQDDRCANLFEALVGTILKAARRKIVTY 61
Db  32 VNLKVRPHSIHALMGENGAGKSTLLKCLFGI-YOKDSGSILFOGKEIDFHSK----- 83

Qy  62 AGELLQG---VHDDVDIVL 78
Db  84 --EALENGISMVHQLNLVL 101

RESULT 5
US-09-756-526A-2
; Sequence 2, Application US/09756526A
; Patent No. US20020038005A1
; GENERAL INFORMATION:
; APPLICANT: Jana, Wojciechowska
; APPLICANT: Evgeny, Lewitin
; APPLICANT: Ludmila, Revina
; APPLICANT: Igor, Zalunin
```

```

Matches 26; Conservative 10; Mismatches 26; Indels 18; Gaps 5;

Qy  7 VNLV--EEIHRIGSKNADGK---LSVKFGVLFQDDRCANLFEALVGTILKAARRKIVTY 61
Db  32 VNLKVRPHSIHALMGENGAGKSTLLKCLFGI-YOKDSGSILFOGKEIDFHSK----- 83

Qy  62 AGELLQG---VHDDVDIVL 78
Db  84 --EALENGISMVHQLNLVL 101

RESULT 4
US-09-815-242-11757
; Sequence 11757, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11757
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae

US-09-815-242-11757

Query Match      14.08; Score 57; DB 10; Length 506;
Best Local Similarity 32.5%; Pred. No. 20;
Matches 26; Conservative 10; Mismatches 26; Indels 18; Gaps 5;

Qy  7 VNLV--EEIHRIGSKNADGK---LSVKFGVLFQDDRCANLFEALVGTILKAARRKIVTY 61
Db  32 VNLKVRPHSIHALMGENGAGKSTLLKCLFGI-YOKDSGSILFOGKEIDFHSK----- 83

Qy  62 AGELLQG---VHDDVDIVL 78
Db  84 --EALENGISMVHQLNLVL 101

RESULT 5
US-09-756-526A-2
; Sequence 2, Application US/09756526A
; Patent No. US20020038005A1
; GENERAL INFORMATION:
; APPLICANT: Jana, Wojciechowska
; APPLICANT: Evgeny, Lewitin
; APPLICANT: Ludmila, Revina
; APPLICANT: Igor, Zalunin
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 02:06:44 ; Search time 71 Seconds
(without alignments)
22.678 Million cell updates/sec

Title: US-09-648-310-2
Perfect score: 406
Sequence: 1 MNVEHVNLLVEIHRGSK.....AGELLQGVHDDVDIVLLQD 81

Scoring table: BLOSOM62
Gapop 10.0 , Gapext 0.5

Searched: 120991 seqs, 19878514 residues

Total number of hits satisfying chosen parameters: 120991

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications-AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	390	96.1	142	10	US-09-925-300-1485 Sequence 1485, Ap
2	61	15.0	323	10	US-09-971-791-6 Sequence 6, Appl
3	57	14.0	506	10	US-09-815-242-5045 Sequence 5045, Ap
4	57	14.0	506	10	US-09-815-242-11757 Sequence 11757, A
5	57	14.0	1163	10	US-09-756-526A-2 Sequence 2, Appl
6	56.5	13.9	474	10	US-09-815-242-10270 Sequence 10270, A
7	56	13.8	506	10	US-09-815-242-10214 Sequence 10214, A
8	55.5	13.7	1160	10	US-09-815-242-10047 Sequence 10047, A
9	55	13.5	1214	10	US-09-783-320-4 Sequence 4, Appl
10	54.5	13.4	239	10	US-09-767-041-10 Sequence 10, Appl
11	54.5	13.4	309	10	US-09-825-882-12 Sequence 12, Appl
12	54.5	13.4	549	10	US-09-115-150-4 Sequence 4, Appl
13	54.5	13.4	639	10	US-09-925-301-1158 Sequence 1158, Ap
14	54.5	13.4	2516	10	US-09-817-514A-2 Sequence 2, Appl
15	54	13.3	859	10	US-09-788-657-16 Sequence 16, Appl
16	53.5	13.2	162	10	US-09-864-761-43711 Sequence 43711, A
17	53.5	13.2	394	9	US-09-738-626-3505 Sequence 3505, Ap
18	53.5	13.2	402	10	US-09-854-731-16 Sequence 16, Appl
19	53	13.1	193	10	US-09-815-242-11663 Sequence 11663, A

20	53	13.1	367	10	US-09-919-497-60 Sequence 60, Appl
21	53	13.1	506	10	US-09-864-761-46512 Sequence 46512, A
22	52.5	12.9	476	10	US-09-779-307-16 Sequence 16, Appl
23	52.5	12.9	477	10	US-09-779-307-4 Sequence 4, Appl
24	52.5	12.9	608	9	US-09-738-626-3609 Sequence 3609, Ap
25	52.5	12.9	689	9	US-10-113-085-2 Sequence 2, Appl
26	52.5	12.9	739	9	US-09-738-626-6988 Sequence 6988, Ap
27	52.5	12.9	771	9	US-09-712-363-157 Sequence 157, App
28	52.5	12.9	872	10	US-09-815-242-13594 Sequence 13594, A
29	52.5	12.9	876	10	US-09-815-242-12623 Sequence 12623, A
30	52	12.8	122	10	US-09-815-242-4935 Sequence 4935, Ap
31	52	12.8	122	10	US-09-815-242-10525 Sequence 10525, A
32	52	12.8	271	10	US-09-815-242-11624 Sequence 11624, A
33	52	12.8	510	9	US-10-051-902-22 Sequence 22, Appl
34	52	12.8	510	9	US-10-051-909-22 Sequence 22, Appl
35	51.5	12.7	272	9	US-09-738-626-5598 Sequence 5598, Ap
36	51.5	12.7	296	10	US-09-843-164-4 Sequence 4, Appl
37	51.5	12.7	426	10	US-09-815-242-5383 Sequence 5383, Ap
38	51.5	12.7	429	10	US-09-815-242-12502 Sequence 12502, A
39	51.5	12.7	751	10	US-09-864-761-38419 Sequence 38419, A
40	51.5	12.7	850	10	US-09-888-615-108 Sequence 108, App
41	51.5	12.7	1175	10	US-09-771-161A-224 Sequence 224, App
42	51.5	12.7	1175	10	US-09-771-161A-225 Sequence 225, App
43	51.5	12.7	1175	10	US-09-771-161A-226 Sequence 226, App
44	51.5	12.7	2025	10	US-09-815-242-5703 Sequence 5703, Ap
45	51.5	12.7	3158	10	US-09-815-242-12611 Sequence 12611, A

ALIGNMENTS

RESULT 1
US-09-925-300-1485
; Sequence 1485, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 1485
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1485

Query Match 96.1%; Score 390; DB 10; Length 142;
Best Local Similarity 93.8%; Pred. No. 4.1e-42;
Matches 76; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MNVEHVNLLVEIHRGSKNADGKLSVKGVLFOODRCANLFEALVGTAKAKRKIVT 60
Db 62 MNVDHEVNLLVEIHRGSKNADGKLSVKGVLFRDDKCANLFEALVGTAKAKRKIVT 121
Qy 61 YAGELLQGVHDDVDIVLLQD 81
Db 122 YPGELLQGVHDDVDIVLLQD 142

RESULT 2
US-09-971-791-6
; Sequence 6, Application US/09971791
; Patent No. US20020115120A1
; GENERAL INFORMATION:
; APPLICANT: Rosanna Kapeller-Libermann